Large scale perturbations in the open universe

David H. Lyth[†] and Andrzej Woszczyna[∗]

[†]*School of Physics and Materials,*
University of Lancaster, Lancaster LA1 4YB.   U. K.

and

Isaac Newton Institute, 20 Clarkson Road, Cambridge CB3 0EH.   U. K.

[∗]*Astronomical Observatory, Jagiellonian University,*
ul. Orla 171, Krakow 30244.   Poland.

Abstract

When considering perturbations in an open ($\Omega_0 < 1$) universe, cosmologists retain only sub-curvature modes (defined as eigenfunctions of the Laplacian whose eigenvalue is less than -1 in units of the curvature scale, in contrast with the super-curvature modes whose eigenvalue is between -1 and 0). Mathematicians have known for almost half a century that all modes must be included to generate the most general *homogeneous Gaussian random field*, despite the fact that any square integrable *function* can be generated using only the sub-curvature modes. The former mathematical object, not the latter, is the relevant one for physical applications. The mathematics is here explained in a language accessible to physicists. Then it is pointed out that if the perturbations originate as a vacuum fluctuation of a scalar field there will be no super-curvature modes in nature. Finally the effect on the cmb of any super-curvature contribution is considered, which generalizes to $\Omega_0 < 1$ the analysis given by Grishchuk and Zeldovich in 1978. A formula is given, which is used to estimate the effect. In contrast with the case $\Omega_0 = 1$, the effect contributes to all multipoles, not just to the quadrupole. It is important to find out whether it has the same l dependence as the data, by evaluating the formula numerically.

1   Introduction

On grounds of simplicity, the present energy density Ω_0 of the universe is generally assumed to be equal to unity (working as usual in units of the critical density).[1] It is not however well determined by observation [1]. The density of baryonic matter can only be of order 0.1 or there will be a conflict with the nucleosynthesis calculation, and although non-baryonic matter seems to be required by observation [2] there is no guarantee that it will bring the total up to $\Omega_0 = 1$. Nor should one assume that a cosmological constant or other exotic contribution to the energy density will play this role.

From a theoretical viewpoint the value $\Omega_0 = 1$ is the most natural, because any other value of Ω is time dependent. The preference for $\Omega_0 = 1$ is sharpened if, as is widely believed, the hot big bang is preceded by an era of inflation. In that case Ω has its present value at the epoch when the present Hubble scale leaves the horizon, and for a generic choice of the inflaton potential this indeed implies that Ω_0 is very close to 1 more or less independently of

---

[1]Throughout this article $\Omega_0 = 1$ will mean a value of Ω_0 close to 1, and $\Omega_0 < 1$ will mean a value substantially less than 1, say less than 0.9.

the initial value of Ω. It is also easier for inflation to explain the homogeneity and isotropy of the observable universe if $\Omega_0 = 1$. On the other hand it is certainly not the case that $\Omega_0 = 1$ is an unambiguous prediction of inflation [3, 4].

The literature on the $\Omega_0 < 1$ cosmology is small compared with the enormous output on the case $\Omega_0 = 1$, because the latter is simpler and observations that can distinguish the two are only now becoming available. This is especially true in regard to the subject of the present paper, which is the effect of spatial curvature on cosmological perturbations. The only data relevant to this subject are the lowest few multipoles of the cosmic microwave background (cmb) anisotropy, that were measured recently by the COBE satellite [5, 6, 7].

This article is concerned both with the basic formalism that one should use in describing cosmological perturbations, and with the cmb multipoles. To describe its contents, let us begin by recalling the presently accepted framework within which cosmological perturbations are discussed.

Cosmological perturbations are expanded in a series of eigenfunctions of the Laplacian for two separate reasons. One is that each mode (each term in the series) evolves independently with time, which makes it easier to evolve a given initial perturbation forward in time. The other is that by assigning a Gaussian probability distribution to the amplitude of each mode, one can generate a homogeneous Gaussian random field. Such a field consists of an ensemble of possible perturbations, and it is supposed that the perturbation seen in the observable universe is a typical member of the ensemble. The stochastic properties of a Gaussian random field are determined by its two point correlation function $\langle f(1)f(2) \rangle$, where f is the perturbation and the brackets denote the ensemble average, and the adjective 'homogeneous' indicates that the correlation function depends only on the distance between the two points.

The question arises which eigenfunctions to use, and in particular what range of eigenvalues to include. If $\Omega_0 = 1$ space is flat and it is known that the Fourier expansion, which includes all negative eigenvalues, is the correct choice. It is complete in two distinct respects. First, it gives the most general *square integrable function*, so that initial conditions in a finite region of the universe can be evolved forward in time. Secondly, it gives the most general *homogeneous Gaussian random field*. Instead of the Fourier expansion one can use the entirely equivalent expansion in spherical polar coordinates.

If $\Omega_0 < 1$, the curvature of space defines a length scale. The spherical coordinate expansion can still be used, and it is known [8, 9] that the modes which have real negative eigenvalue *less than* -1 *in units of the curvature scale* provide a complete orthonormal basis for square integrable functions. Presumably for this reason, only these modes have been retained by cosmologists. We will call them *sub-curvature* modes, because they vary significantly on a scale which is less than the curvature scale. The other modes, with eigenvalues between -1 and 0 in units of the curvature scale, we will call super-curvature modes.

It is certainly enough to retain only sub-curvature modes if all one wishes to do is to track the evolution of a given initial perturbation, since the region of interest is always going to be finite and any function defined in a finite region can be expanded in terms of the sub-curvature modes. (In fact, to describe the observations that we can make it is enough to specify initial conditions within our past light cone.) But this is not what one does in cosmology.[2] Rather, one uses the mode expansion to generated a Gaussian perturbation, by assigning a Gaussian probability distribution to the amplitude of each mode. In this context the inclusion of only sub-curvature modes looks restrictive. For example, it leads to a correlation function which necessarily becomes small at distances much bigger than the

---

[2]The only case where one is interested in evolving a given initial condition is when one obtains the well known relation between the matter density contrast and the galaxy peculiar velocity field, but one uses this prediction only on very small scales where the curvature cannot be significant. Even then, one is still interested in the stochastic properties as well.

curvature scale (to be precise, it is less than $r/\sinh r$ times its value at $r = 0$, where r is the distance in curvature units).

Faced with this situation, we queried the assumption that only sub-curvature modes should be included, and the results of our investigation are reported here.

First we describe the mathematical situation, showing that indeed a more general Gaussian random field is generated by including also the super-curvature modes. As expected the correlation function can now be constant out to arbitrarily large distances.

Then we go on to ask whether nature has chosen to use the super-curvature modes, focussing on the low multipoles of the cmb anisotropy which are the only relevant observational data, and on the curvature perturbation which is thought to be responsible for these multipoles. If, as is usually supposed, this perturbation originates as a vacuum fluctuation of the inflaton field, there will be no super-curvature modes. On the other hand, like any other statement about the universe one expects this assumption to be at best approximately valid. Supposing that it fails badly on some very large scale, *but that the curvature perturbation still corresponds to a typical realization of a homogeneous Gaussian random field*, one is lead to ask if a failure of the assumption could be detected by observing the cmb anisotropy. We note that for $\Omega_0 = 1$ this question has already been discussed by Grishchuk and Zeldovich [10], and we extend their discussion to the case $\Omega_0 < 1$.

After our investigation was complete, and the draft of this paper was almost complete, M. Sasaki suggested to one of us (DHL) that a mathematics paper written by Yaglom in 1961 [11] might be relevant. From this paper we learned that the need to include both sub- and super-curvature modes in the expansion of a homogeneous Gaussian random field in negatively curved space has been known to mathematicians since at least 1949 [12]. It would appear therefore that the assumption by cosmologists that only the sub-curvature modes are needed is a result of a complete failure of communication between the worlds of mathematics and science, which has persisted for many decades. We have retained the mathematics part of our paper because it gives the relevant results in the sort of language that is familiar to physicists, though it is strictly speaking redundant.

Let us end this introduction by saying a bit more about the cosmology literature. Starting with the paper of Lifshitz in 1946 [13], there are many papers on the treatment of cosmological perturbations for the case $\Omega < 1$. However, most of them deal with the *definition* and *evolution* of the perturbations, which is not our main concern. We have not attempted a full survey of this part of the literature, but have just cited useful papers that we happen to be aware of. By contrast, the cosmology literature on *stochastic properties* is very small for the case $\Omega_0 < 1$, and as we have mentioned it is out of touch with the relevant pure mathematics literature where the theory of random fields is discussed. The first serious treatment of stochastic properties is by Wilson in 1983 [14]. He developed the theory from scratch, and not surprisingly included only the sub-curvature modes which he knew were sufficient for the description of the non-stochastic properties. His notation is defective and much is left unsaid, but subsequent papers have not made basic advances in the formulation of the subject, though they have gone much further in calculating the cmb multipoles and comparing them with observation. We believe our referencing to be reasonable complete, as far as the cosmology literature on the stochastic properties is concerned.

The layout of this paper is as follows. In Section 2 some basic formulas are given for the Robertson-Walker universe with $\Omega < 1$. In Section 3 the standard procedure is described, and in the next section it is extended to the super-curvature modes. Inflation is discussed in Section 5, and the cmb anisotropy is treated in Section 6. In an Appendix we give various mathematical results in the sort of language that is familiar to us as physicists.

2   Distance scales

Ignoring perturbations, the universe is homogeneous and isotropic. There is a universal scale factor $a(t)$, with t the universal time measured by the synchronized clocks of comoving observers, and the distance between any two such observers is proportional to a.

According to the Einstein field equation, the time dependence of a is governed by the Friedmann equation which may be written

$$1 - \Omega = -\frac{K}{(aH)^2} \tag{1}$$

Here K is a constant, $H = \dot{a}/a$ is the Hubble parameter, and Ω is the energy density measured in units of the critical density $3H^2/8\pi G$. (As usual we set $c = 1$, and regard any cosmological constant as a contribution to the energy density as opposed to a modification of the Einstein field equation.) The spatial curvature scalar is

$$R^{(3)} = 6K/a^2 \tag{2}$$

The distance $a/|K|$ defines the curvature scale; on much smaller scales space is practically flat, whereas on much bigger scales the effect of curvature is very important. From Eq. (1) the Hubble distance $1/H$ is a fraction $(1 - \Omega)^{1/2} < 1$ of the curvature scale. Even in the extreme case $\Omega_0 = 0.1$ this makes the curvature scale about three times the present Hubble distance.

We will set $K = -1$ so that a is the curvature scale. Then the case $\Omega = 1$ corresponds to the limit $a \to \infty$, with physical distances like H^{-1} remaining constant. Note that the effect of curvature in a comoving region becomes neither more nor less important with the passage of time, since the curvature scale expands with the universe.

We are concerned with the comoving region which is now the observable universe, bounded by the surface of last scattering of radiation emitted at very high redshift. This is close to the particle horizon of a matter dominated cosmology, unless there is a cosmological constant or some other non-standard contribution to the energy density. The coordinate distance of this particle horizon [15] (ie., its distance in units of the curvature scale) is r_{ph} where $\sinh^2 \frac{1}{2} r_{\mathrm{ph}} = \Omega_0^{-1} - 1$. Even the smallest conceivable value $\Omega_0 \simeq 0.1$ gives $r_{\mathrm{ph}} = 3.6$, so effect of curvature is negligible except on scales comparable with the size of the observable universe.

From Eq. (1), the physical distance of the particle horizon is

$$a_0 r_{\mathrm{ph}} = (1 - \Omega_0)^{-1/2} H_0^{-1} r_{\mathrm{ph}} \tag{3}$$

For $\Omega_0 = 1$ it is $2H_0^{-1}$, and even for $\Omega_0 = 0.1$ it is only $3.8H_0^{-1}$. Thus it is not very much bigger than the Hubble distance H_0^{-1}.

3   Sub-curvature modes

We are concerned with the first order treatment of cosmological perturbations. To this order, the perturbations 'live' in unperturbed spacetime, because the distortion of the spacetime geometry is itself a perturbation.

The perturbations satisfy linear partial differential equations, in which derivatives with respect to comoving coordinates occur only through the Laplacian. When the perturbations are expanded in eigenfunctions of the Laplacian with eigenvalues $-(k/a)^2$, each mode (term in the expansion) decouples.

Denoting the eigenvalue by $-(k/a)^2$, it is known [8, 9] that the modes with real $k^2 > 1$ provide a complete orthonormal basis for L^2 functions, and the usual procedure is to keep

only them. Since they all vary appreciably on scales less than the curvature scale a we will call them sub-curvature modes. It will be useful to define the quantity

$$q^2 = k^2 - 1 \tag{4}$$

3.1 The spherical expansion

Spherical coordinates are defined by the line element

$$dl^2 = a^2[dr^2 + \sinh^2 r(d\theta^2 + \sin^2\theta d\phi^2)] \tag{5}$$

In the region $r \ll 1$ curvature is negligible and this becomes the flat-space line element written in spherical polar coordinates. The volume element between adjacent spheres is $4\pi \sinh^2 r dr$, so for $r \gg 1$ the volume $\mathcal{V}$ and area $\mathcal{A}$ of a sphere are related by $\mathcal{V} = \mathcal{A}/2$. In contrast with the flat-space case this relation is independent of r, because most of the volume of a very large sphere is near its surface.

Since the spherical harmonics Y_{lm} are a complete set on the sphere, any eigenfunction can be expanded in terms of them. The radial functions depend only on r, and they satisfy a second order differential equation. As in the flat-space case, only one of the two solutions is well behaved at the origin, so the radial functions are completely determined up to normalisation. The mode expansion of a generic perturbation f is therefore of the form

$$f(r, \theta, \phi, t) = \int_0^\infty dq \sum_{lm} f_{klm}(t) Z_{klm}(r, \theta, \phi) \tag{6}$$

where

$$Z_{klm} = \Pi_{kl}(r) Y_{lm}(\theta, \phi) \tag{7}$$

A compact expression for the radial functions is [16, 13, 17, 9, 18]

$$\Pi_{kl} = \frac{\Gamma(l + 1 + iq)}{\Gamma(iq)} \sqrt{\frac{1}{\sinh r}} P_{iq-\frac{1}{2}}^{-l-\frac{1}{2}}(\cosh r) \tag{8}$$

which corresponds to the normalisation

$$\int_0^\infty \Pi_{kl}(r) \Pi_{k'l'}(r) \sinh^2 r dr = \delta(q - q')\delta_{ll'} \tag{9}$$

The corresponding normalisation of the eigenfunctions is

$$\int Z_{klm}^* Z_{k'l'm'} d\mathcal{V} = \delta(q - q')\delta_{ll'}\delta_{mm'} \tag{10}$$

where

$$d\mathcal{V} = \sinh^2 r \sin\theta dr d\theta d\varphi \tag{11}$$

is the volume element.

As it stands Eq. (8) has a constant nonzero phase. It is convenient to drop this phase so that the function is real, and one then has the explicit expressions [18][3]

$$\Pi_{kl} \equiv N_{kl}\tilde{\Pi}_{kl} \tag{12}$$

$$\tilde{\Pi}_{kl} \equiv q^{-2}(\sinh r)^l \left(\frac{-1}{\sinh r}\frac{d}{dr}\right)^{l+1} \cos(qr) \tag{13}$$

$$N_{kl} \equiv \sqrt{\frac{2}{\pi}}q^2 \left[\prod_{n=0}^l (n^2 + q^2)\right]^{-1/2} \tag{14}$$

[3]These expressions correct some misprints in [19, 4].

The un-normalised radial functions $\tilde{\Pi}_{kl}$ satisfy a recurrence relation [20]

$$\tilde{\Pi}_{k,l+2} = -\left[(l+1)^2 + q^2\right]\tilde{\Pi}_{kl} + (2l+3)\coth r\,\tilde{\Pi}_{k,l+1} \tag{15}$$

and the first three functions are

$$\tilde{\Pi}_{k0} = \frac{1}{\sinh r}\left[\frac{\sin(qr)}{q}\right] \tag{16}$$

$$\tilde{\Pi}_{k1} = \frac{1}{\sinh r}\left[-\cos(qr) + \coth r\frac{\sin(qr)}{q}\right] \tag{17}$$

$$\tilde{\Pi}_{k2} = \frac{1}{\sinh r}\left[-3\coth r\cos(qr) + (3\coth^2 r - q^2 - 1)\frac{\sin(qr)}{q}\right] \tag{18}$$

The case $\Omega = 1$ corresponds to $q \to \infty$ with qr fixed, and in that limit $\Pi_{kl}(r)$ reduces to the familiar radial function,

$$\Pi_{kl}(r) \to \sqrt{\frac{2}{\pi}}qj_l(qr). \tag{19}$$

Near the origin $\Pi_{kl}(r)$ has the same behaviour as $j_l(qr)$, namely $\Pi_{kl} \propto r^l$, which ensures that the Laplacian is well defined there. The other linearly independent solution of the radial equation, which corresponds to the substitution $\cos(qr) \to \sin(qr)$ in Eq. (13), has the same behaviour as the other Bessel function $h_l(qr)$ and is therefore excluded.

3.2 Stochastic properties

We are interested in the stochastic properties of the perturbations, at fixed time. To define them we will take the approach of considering an ensemble of universes of which ours is supposed to be one.

The stochastic properties of a generic perturbation $f(r,\theta,\phi)$ are defined by the set of probability distribution functions, relating to the outcome of a simultaneous measurement of a perturbation at a given set of points. From the probability distributions one can calculate ensemble expectation values, such as the correlation function for a pair of points r_1,θ_1,ϕ_1 and r_2,θ_2,ϕ_2,

$$\xi_f \equiv \langle f(r_1,\theta_1,\phi_1), f(r_2,\theta_2,\phi_2)\rangle \tag{20}$$

and the mean square $\langle f^2(r,\theta,\phi)\rangle$.

If the probability distributions depend only on the geodesic distances between the points, the perturbation is said to be *homogeneous* with respect to the group of transformations that preserve this distance. (For flat space this is the group of translations and rotations, and for homogeneous negatively curved space it is isomorphic to the Lorentz group [21].) Then the correlation function depends only on the distance between the points, and the mean square is just a number.

Cosmological perturbations are assumed to be homogeneous, and except for the curvature perturbation that we discuss in Section 6 their correlation functions are supposed to be very small beyond some maximum distance, called the correlation length.

An ergodic universe?

If there is a finite correlation length, one ought to be able to dispense with the concept of an ensemble of universes, in favour of the concept of sampling our own universe at different locations. In this approach one defines the probability distribution for simultaneous measurements at N points with by considering random locations of these points, subject to the condition that the distances between them are fixed. The correlation function is

defined by averaging over all pairs of points a given distance apart, and the mean square is the spatial average of the square. For a Gaussian perturbation in flat space this 'ergodic' property can be proved under weak conditions [22] and there is no reason to think that spatial curvature causes any problem though we are not aware of any literature on the subject.

For the ergodic viewpoint to be useful, the observable in question has to be measured in a region that is big compared with the correlation length. This is the case for the distributions and peculiar velocities of galaxies and clusters, where surveys have been done out to several hundred Mpc to be compared with a correlation length of order 10 Mpc, and accordingly the ergodic viewpoint is always adopted there [23]. However, even a distance of a few hundred Mpc is only ten percent or so of the Hubble distance H_0^{-1}, and therefore at most a few percent of the curvature scale $(1 - \Omega_0)^{-1/2} H_0^{-1}$. Thus galaxy and cluster surveys do not probe spatial curvature. The only observables that do, which are the low multipoles of the cmb anisotropy, are measured only at our position so there is no practical advantage in going beyond the concept of the ensemble even if the mathematics turns out to be straightforward.

In addition to the interpretation that the ensemble corresponds to different locations within the smooth patch of the universe that we inhabit, there are two other possibilities. One is that the ensemble corresponds to different smooth patches, which are indeed supposed to exist both in 'chaotic' [24] and bubble nucleation [25, 26, 27, 28] scenarios of inflation. The other, adopting the usual language of quantum mechanics, is to regard the ensemble as the set of all possible outcomes of a 'measurement' performed on a given state vector. A concrete realization of this 'quantum cosmology' viewpoint is provided by the hypothesis that the perturbations originate as a vacuum fluctuation of the inflaton field, which we consider later.

3.3 Gaussian perturbations

It is generally assumed that cosmological perturbations are Gaussian, in the regime where they are evolving linearly. A Gaussian perturbation is normally defined as one whose probability distribution functions are multivariate Gaussians [29, 22, 30], and its stochastic properties are completely determined by its correlation function. The perturbation is homogeneous if the correlation function depends only on the distance between the points.

The simplest Gaussian perturbation is just a coefficient times a given function, the coefficient having a Gaussian probability distribution. A more general Gaussian perturbation is a linear superposition of functions [29],

$$f(r, \theta, \phi) = \sum_n f_n X_n(r, \theta, \phi) \tag{21}$$

with each coefficient having an independent Gaussian distribution. Its stochastic properties are completely determined by the mean squares $\langle f_n^2 \rangle$ of the coefficients. (For the moment we are taking the expansion functions X_n to be real, and to be labelled by a discrete index.)

The correlation function corresponding to the above expansion is

$$\langle f(r_1, \theta_1, \phi_1) f(r_2, \theta_2, \phi_2) \rangle = \sum_n \langle f_n^2 \rangle X_n(r_1, \theta_1, \phi_1) X_n(r_2, \theta_2, \phi_2) \tag{22}$$

For it to depend only on the distance between the points requires very special choices of the expansion functions, and of the mean squares $\langle f_n^2 \rangle$.

It is very important to realise that the functions in such an expansion need not be linearly independent. Suppose for example that $X_3 = X_1 + X_2$, and that $\langle f_3^2 \rangle$ is much bigger than $\langle f_1^2 \rangle$ and $\langle f_2^2 \rangle$. Then most members of the ensemble are of the form $f = \text{const} X_3$, which would clearly not have been the case if the function X_3 had been dropped because of its linear dependence.

So far all our considerations have been at a fixed time. The time dependence is trivial if we expand in eigenfunctions of the Laplacian, because each coefficient f_n then evolves independently of the others. Let us therefore replace the discrete, real expansion above by the complex, partially continuous expansion Eq. (6). The coefficients now satisfy the reality condition $f^*_{klm} = f_{kl-m}$, and a Gaussian perturbation is constructed by assigning independent Gaussian probability distributions to the real and imaginary parts of the coefficients with $m \geq 0$. We demonstrate in the Appendix that the correlation function being dependent only on the distance between the points is equivalent to the mean squares of their real and imaginary parts being equal, and independent of l and m. One can therefore define the *spectrum* of a generic perturbation f by [4]

$$\langle f^*_{klm} f_{k'l'm'} \rangle = \frac{2\pi^2}{q(q^2+1)} \mathcal{P}_f(k) \delta(q-q') \delta_{ll'} \delta_{mm'} \tag{23}$$

The q dependence of this expression has been chosen to give the simple form Eq. (30) for the mean square perturbation. In the flat-space limit $q \to \infty$ it reduces to

$$\langle f^*_{klm} f_{k'l'm'} \rangle = \frac{2\pi^2}{k^3} \mathcal{P}_f(k) \delta(k-k') \delta_{ll'} \delta_{mm'} \tag{24}$$

The correlation function is given by

$$\xi_f = \int_0^\infty dq \frac{2\pi^2}{q(q^2+1)} \mathcal{P}_f(k) \sum_{lm} Z^*_{klm}(r_1, \theta_1, \phi_1) Z_{klm}(r_2, \theta_2, \phi_2) \tag{25}$$

Taking one of the points to be at the origin, only the term $l = m = 0$ survives, and using

$$Z_{k00}(r, \theta, \phi) = (2\pi^2)^{-1/2} q \sin(qr)/(q \sinh r) \tag{26}$$

one finds

$$\xi_f(r) = \int_1^\infty \frac{dk}{k} \mathcal{P}_f(k) \frac{\sin(qr)}{q \sinh r} \tag{27}$$

Note the appearance in this expression of the logarithmic interval

$$\frac{dk}{k} = \frac{q dq}{1+q^2} \tag{28}$$

The flat-space limit is $q \to \infty$ with qr fixed, and the correlation function then reduces to

$$\xi_f(r) = \int_0^\infty \mathcal{P}_f(k) \frac{\sin(kr)}{kr} \frac{dk}{k} \tag{29}$$

(There is no distinction between k and q in the flat-space limit, and whenever we consider that limit we will use the the symbol k.) Setting $r = 0$ gives the mean square value

$$\xi_f(0) \equiv \langle f^2 \rangle = \int_1^\infty \frac{dk}{k} \mathcal{P}_f(k) \tag{30}$$

The flat-space limit is

$$\xi_f(0) \equiv \langle f^2 \rangle = \int_0^\infty \frac{dk}{k} \mathcal{P}_f(k) \tag{31}$$

Since the spectrum is positive and $|\sin(qr)| < qr$, the flat-space correlation function is never bigger than its value at $r = 0$, but without knowing the spectrum one cannot say more. The situation is very different for the curved-space expression Eq. (27), because it contains an extra factor $r/\sinh r$. It follows that by expanding a perturbation in terms of sub-curvature modes one obtains a correlation function bounded by

$$\frac{\xi_f(r)}{\xi_f(0)} < \frac{r}{\sinh r} \tag{32}$$

In order for $\langle f^2 \rangle$ to be well defined, the spectrum must have appropriate behaviour at $q = \infty$ and 0. As $q \to \infty$ one needs $\mathcal{P} \to 0$. As $q \to 0$ one needs $\mathcal{P} \to 0$ in the flat case, but only $q^2 \mathcal{P}_f(k) \to 0$ in the curved case.

Note that in the curved case the limit $q \to 0$ does *not* correspond to infinite large scales, but rather to scales of order the curvature scale. This means that one cannot tolerate a divergent behaviour there (unless of course the curvature scale happens to be larger than any relevant scale, in which case we are back to flat space).

For future reference, we note that most other authors have used a different definition of the spectrum. This is usually denoted by P_f, and it is related to our $\mathcal{P}_f$ by

$$\mathcal{P}_f(k) = \frac{q(q^2 + 1)}{2\pi^2} P_f(k) \tag{33}$$

With this definition,

$$\xi_f(r) = \frac{1}{2\pi^2} \int_0^\infty \mathrm{d}q q^2 P_f(k) \frac{\sin(qr)}{q \sinh r} \tag{34}$$

4   Including the super-curvature modes

In the last section we found that the usual procedure, which includes only sub-curvature modes, generates a Gaussian perturbation whose correlation function necessarily falls off faster than $r/\sinh r$. This reflects the fact that each super-curvature mode varies strongly on a scale no bigger than the curvature scale. A random superposition of such modes will hardly ever be nearly constant on a scale much bigger than the curvature scale [31], which is precisely what the lack of correlation on large scales is telling us.

Faced with this situation one can might think that the lack of correlation on super-curvature scales is just a mathematical fact, inherent in the nature of homogeneous negatively curved space. Certainly it is not trivial to construct a function exhibiting super-curvature correlations. Consider, for instance, the following construction; throw down randomly spheres with radius much bigger than the curvature scale and fill them uniformly with galaxies, leaving the rest of the universe empty. Then one might think that a typical observer will see a uniform distribution of galaxies out to a distance much bigger than the curvature scale, making the correlation function almost flat out to such a distance. But this is incorrect, because according to the line element Eq. (5) most of the volume of a sphere is near its edge, and so is a typical observer.[4]

This example notwithstanding, correlation on arbitrarily large scales *is* possible, and is achieved simply by including the super-curvature modes.

For $-1 < q^2 < 0$ the analytic continuation of the radial function Π_{kl} is purely imaginary, and for convenience we drop the i factor. Thus the super-curvature modes are defined by

$$\Pi_{kl} \equiv N_{kl}\tilde{\Pi}_{kl} \tag{35}$$

$$\tilde{\Pi}_{kl} \equiv |q|^{-2}(\sinh r)^l \left(\frac{-1}{\sinh r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^{l+1} \cosh(|q|r) \tag{36}$$

$$N_{k0} \equiv \sqrt{\frac{2}{\pi}}|q| \tag{37}$$

$$N_{kl} \equiv \sqrt{\frac{2}{\pi}}|q| \left[\prod_{n=1}^{l}(q^2 + n^2)\right]^{-1/2} \qquad (l > 0) \tag{38}$$

The recurrence relation Eq. (15) is still satisfied, and the first three functions are

$$\tilde{\Pi}_{k0} = \frac{1}{\sinh r}\left[\frac{\sinh(|q|r)}{|q|}\right] \tag{39}$$

---

[4]One of us (DHL) is indebted to R. Gott and P. J. E. Peebles for pointing out this fact.

$$\tilde{\Pi}_{k1} = \frac{1}{\sinh r}\left[-\cosh(|q|r) + \coth r\frac{\sinh(|q|r)}{|q|}\right] \tag{40}$$

$$\tilde{\Pi}_{k2} = \frac{1}{\sinh r}\left[-3\coth r\cosh(|q|r) + (3\coth^2 r - q^2 - 1)\frac{\sinh(|q|r)}{|q|}\right] \tag{41}$$

At large r the super-curvature modes go like $\exp[-(1-|q|)r]$. Because the volume element is $d\mathcal{V} = \sinh^2 r\sin\theta dr d\theta d\varphi$ the integral over all space of a product of any two of them diverges. As a result they are not orthogonal in the sense of Eq. (9), let alone orthonormal. In any finite region of space (and of course we are only going to do physics in such a region) they are not even linearly independent of the sub-curvature eigenfunctions, since the latter are complete (for the set of L^2 functions defined over all space). None of this matters for the purpose of generating a Gaussian perturbation.

The super-curvature modes add an additional term to the expansion Eq. (6),

$$f^{\rm SC}(r,\theta,\phi) = \int_0^1 d(iq)\sum_{lm} f_{klm}Z_{klm}(r,\theta,\phi) \tag{42}$$

Let us define the corresponding spectrum by analogy with Eq. (23),

$$\langle f_{klm}f_{k'l'm'}^*\rangle = \frac{2\pi^2}{|q|(q^2+1)}\mathcal{P}_f(k)\delta(|q|-|q'|)\delta_{ll'}\delta_{mm'} \quad (-1 < q^2 < 0) \tag{43}$$

We show in the Appendix that the correlation function remains well defined, and dependent only on the distance between the two points. Taking one of them to be at the origin only the $l = 0$ mode survives and the super-curvature contribution to the correlation function is seen to be

$$\xi_f^{\rm SC}(r) = \int_0^1 \frac{dk}{k}\mathcal{P}_f(k)\frac{\sinh(|q|r)}{|q|\sinh r} \tag{44}$$

The super-curvature contribution to the mean square is

$$\langle f^2\rangle^{\rm SC} = \int_0^1 \frac{dk}{k}\mathcal{P}_f(k) \tag{45}$$

Unified expressions including all modes

The use of q in the mode expansion Eq. (6) is natural for the sub-curvature modes, and we are using in this paper to facilitate comparison with existing literature. Unified expressions including all modes on an equal footing would use k in the mode expansion, so defining new coefficients $\tilde{f}_{klm}$. One would then have the following expressions, which include both sub- and super-curvature modes.

$$f(r,\theta,\phi,t) = \int_0^\infty dk\sum_{lm}\tilde{f}_{klm}(t)Z_{klm}(r,\theta,\phi) \tag{46}$$

$$\langle\tilde{f}_{klm}^*\tilde{f}_{k'l'm'}\rangle = \frac{2\pi^2}{k|q^2|}\mathcal{P}_f(k)\delta(k-k')\delta_{ll'}\delta_{mm'} \tag{47}$$

$$\xi_f(r) = \int_0^\infty \frac{dk}{k}\mathcal{P}_f(k)\frac{\sin(qr)}{q\sinh r} \tag{48}$$

4.1 Very large super-curvature scales

The contribution to the correlation function from a mode with $k^2 \equiv 1 + q^2 \ll 1$ is

$$\xi_f(r) \propto \exp(-k^2 r) \tag{49}$$

for $r \gg 1$. Thus the correlation length, in units of the curvature scale a, is of order k^{-2}. This is in contrast with the flat-space case, where the contribution from a mode with $k \ll 1$ gives a correlation length of order $1/k$. The difference can be understood in terms of the different behaviour of the volume element, in the following way. In both cases, the r dependence is that of the $l = 0$ mode, and as long as r is small enough that the mode is approximately constant the divergence theorem gives

$$\frac{r}{f}\frac{\mathrm{d}f}{\mathrm{d}r} \simeq \frac{-k^2 r \mathcal{V}(r)}{\mathcal{A}(r)} \tag{50}$$

where $\mathcal{V}$ is the volume within a sphere of radius r and $\mathcal{A}$ is the area of this sphere. In flat space the right hand side is equal to $-(kr)^2/3$, so it is small out to a distance $r \sim 1/k$. In curved space, when $r \gg 1$, the line element Eq. (5) shows that most of the volume of the sphere is near its edge, and the right hand side becomes equal to $-k^2 r/2$, which is small out to a distance $r \sim k^{-2}$.

In addition to being significant in its own right, the correlation function determines other physically significant quantities. One, which is relevant for quantities like the density perturbation, is the $\sigma_f^2(r)$, the mean square of f after it has been smeared over a sphere of radius r, which is given by

$$\sigma_f^2(r) = \mathcal{V}(r)^{-2} \int \mathrm{d}\mathcal{V}_1 \int \mathrm{d}\mathcal{V}_2 \xi_f(r_1, \theta_1, \phi_1, r_2, \theta_2, \phi_2) \tag{51}$$

where the integrations are within the spheres $r_1 < r$ and $r_2 < r$ (we have taken advantage of homogeneity to evaluate locate the sphere at the origin). Because most of the volume of the sphere is near it's edge, the behaviour Eq. (49) of the correlation function leads to the same behaviour for $\sigma_f^2(r)$, so it too remains constant out to a distance $r \sim k^{-2}$. Another quantity, which is relevant for the curvature perturbation $\mathcal{R}$ that we shall discuss in Section 6, is the mean square after smearing over a geodesic surface of radius r,

$$\mathcal{A}(r)^{-2} \int \mathrm{d}\mathcal{A}_1 \int \mathrm{d}\mathcal{A}_2 \xi_{\mathcal{R}}(r_1, \theta_1, \phi_1, r_2, \theta_2, \phi_2) \tag{52}$$

On super-curvature scales it is a measure of the fractional perturbation in the curvature of the surface. One sees that it too shares the behaviour Eq. (49), and so is constant out to a distance $r \sim k^{-2}$.

5 The cmb anisotropy

Even if $\Omega_0 < 1$, spatial curvature is negligible on scales that are small compared with the Hubble distance. As a result, the only observational data that can be sensitive to curvature, even if $\Omega_0 < 1$, are the lowest few multipoles of the cmb anisotropy. Papers discussing the effect of curvature on these multipoles [31, 14, 32, 4, 33, 34] appeared sporadically before they were measured by the COBE satellite [5], and many have appeared since [35, 36, 37, 38, 39, 40, 41]. All of these papers keep only sub-curvature modes. Here we consider both sub- and super-curvature modes.

The multipoles are defined by

$$\frac{\Delta T(\mathbf{e})}{T} = \mathbf{w}.\mathbf{e} + \sum_{l=2}^{\infty} \sum_{m=-l}^{+l} a_{lm} Y_{lm}(\mathbf{e}) . \tag{53}$$

The dipole term $\mathbf{w}.\mathbf{e}$ is well measured, and is the Doppler shift caused by our velocity $\mathbf{w}$ relative to the rest frame of the cmb. Unless otherwise stated, ΔT will denote only the intrinsic, non-dipole contribution from now on.

If the perturbations in the universe are Gaussian, the real and imaginary part of each multipole will have an independent Gaussian probability distribution (subject to the condition $a_{lm}^* = a_{l,-m}$). The expectation values of the squares of the real and imaginary parts are equal so one need only consider their sum,

$$C_l \equiv \left\langle |a_{lm}|^2 \right\rangle . \tag{54}$$

Rotational invariance is equivalent to the independence of this expression on m.

Even if it can be identified with an average over observer positions, the expectation value C_l cannot be measured. Given a theoretical prediction for C_l, the best guess for $|a_{lm}|^2$ measured at our position is that it is equal to C_l, but one can also calculate the variance of this guess, which is called the *cosmic variance*. Since the real and imaginary part of each multipole has an independent Gaussian distribution the cosmic variance of $\sum_m |a_{lm}|^2$ is only $2/(2l+1)$ times its expected value, and by taking the average over several l's one can reduce the cosmic variance even further. Nevertheless, for the low multipoles that are sensitive to curvature it represents a serious limitation on our ability to distinguish between different hypotheses about the C_l. Any hypothesis can be made consistent with observation by supposing that the region around us is sufficiently atypical.

The surface of last scattering of the cmb is practically at the particle horizon, whose coordinate distance is η_0 with $\sinh^2 \eta_0/2 = \Omega_0^{-1} - 1$. An angle θ subtends at this surface a coordinate distance d given by [23]

$$\theta = \frac{1}{2}(1 - \Omega_0)^{-1/2}\Omega_0 d = \frac{1}{2}(a_0 H_0 \Omega_0 d) \tag{55}$$

Spatial curvature is negligible when $d \ll 1$, corresponding to

$$\theta \ll 30(1 - \Omega_0)^{-1/2}\Omega_0 \text{ degrees} \tag{56}$$

A structure with angular size θ radians is dominated by multipoles with

$$l \sim 1/\theta \tag{57}$$

one expects that spatial curvature will be negligible for the multipoles

$$l \gg \frac{2\sqrt{1 - \Omega_0}}{\Omega_0} \tag{58}$$

This is the regime $l \gg 20$ if $\Omega_0 = 0.1$, and the regime $l \gg 6$ if $\Omega_0 = 0.3$.

This restriction need not apply to super-curvature modes with $k^2 \ll 1$ because the spatial gradient involved is then small in units of the curvature scale. The contribution of these modes is called the Grishchuk-Zeldovich effect, and we discuss it later.

The linear scale probed by the multipoles decreases as l increases, and for $l \sim 1000$ it becomes of order $100\,\mathrm{Mpc}$. On these scales one can observe the distribution and motion of galaxies and clusters in the region around us. On the supposition that they all have a common origin, the cmb anisotropy and the motion and distribution of galaxies and clusters are collectively termed 'large scale structure'.

A promising model of large scale structure is that it originates as an adiabatic density perturbation, or equivalently [42, 43, 44, 45] as a perturbation in the curvature of the hypersurfaces orthogonal to the comoving worldlines. This model has has been widely investigated for the case $\Omega_0 = 1$ [46], and recently it has been advocated also for the case $\Omega_0 < 1$ [35, 39, 37]. In this paper we consider the model only in relation to the cmb anisotropy since the galaxy and cluster data are insensitive to spatial curvature. We note though that the full data set may impose a significant lower bound on Ω_0 [47].

5.1 The curvature perturbation

The curvature perturbation is conveniently characterised by a quantity $\mathcal{R}$, which is defined in terms of the perturbation in the curvature scalar by[5]

$$4(k^2 + 3)\mathcal{R}_{klm}/a^2 = \delta R^{(3)}_{klm} \tag{59}$$

In the limit $\Omega \to 1$,

$$4k^2 \mathcal{R}_{klm}/a^2 = \delta R^{(3)}_{klm} \tag{60}$$

On cosmologically interesting scales, $\mathcal{R}_{klm}$ is expected to be practically constant in the early universe. To be precise, it is practically constant on scales far outside the horizon in the regime where $\Omega(t)$ is close to 1 (assuming that the density perturbation is adiabatic) [43, 44, 45, 52]. During matter domination the former condition can be dropped, so that $\mathcal{R}_{klm}$ is constant on all scales until Ω breaks away from 1. After that it has the time dependence $\mathcal{R}_{klm} = F\hat{\mathcal{R}}_{klm}$ where $\hat{\mathcal{R}}_{klm}$ is the early time constant value and

$$F = 5\frac{\sinh^2 \eta - 3\eta \sinh \eta + 4\cosh \eta - 4}{(\cosh \eta - 1)^3} \tag{61}$$

with

$$\eta = 2(aH)^{-1} = 2(1 - \Omega)^{1/2} \tag{62}$$

Unless otherwise indicated, $\mathcal{R}$ will indicate the primordial value $\hat{\mathcal{R}}$ from now on.

During matter domination and before Ω breaks away from 1, the density contrast is given by

$$\left(\frac{a^2 H^2}{k^2 + 3}\right)\frac{\delta\rho_{klm}}{\rho} = \frac{2}{5}\mathcal{R}_{klm} \tag{63}$$

For $\Omega_0 = 1$ this reduces to

$$\left(\frac{a^2 H^2}{k^2}\right)\frac{\delta\rho_{klm}}{\rho} = \frac{2}{5}\mathcal{R}_{klm} \tag{64}$$

In these expressions the density perturbation is evaluated on comoving hypersurfaces (it is often referred to as the 'gauge invariant' density perturbation). In the matter dominated era where they hold this is the same as the density perturbation in the 'synchronous gauge' with the 'gauge mode' dropped [53].

5.2 The Sachs-Wolfe effect

Horizon entry occurs long after matter domination on scales

$$a_0/k \gg 20(\Omega_0 h^2)^{-1} \, \text{Mpc} \tag{65}$$

where h is the value of H_0 in units of $100 \, \text{km} \, \text{sec}^{-1} \, \text{Mpc}^{-1}$. As a fraction of the Hubble distance these are the scales

$$\frac{a_0 H_0}{k} \gg \frac{.007}{\Omega_0 h} \tag{66}$$

It follows from Eq. (55) that they correspond to multipoles

$$l \ll 300h \tag{67}$$

---

[5]The quantity $\mathcal{R}$ was called ϕ_m by Bardeen who first considered it [42], $\mathcal{R}_m$ by Kodama and Sasaki [48]. It is equal to $3/2$ times the quantity $\delta K/k^2$ of Lyth [43, 44], which is in turn equal to the ζ of Mukhanov, Feldman and Brandenberger [49]. After matter domination it is equal to $-(3/5)\Phi$, where Φ is the peculiar gravitational potential (and one of the 'gauge invariant' variables introduced in [42]). On scales far outside the horizon, in the case $\Omega = 1$, it is the ζ of [50], and three times the ζ of [51].

Assuming an initial adiabatic perturbation, these multipoles are dominated by the effect of the distortion of the spacetime metric between us and the surface of last scattering, which is called Sachs-Wolfe effect. If $\Omega_0 = 1$ the Sachs-Wolfe approximation accounts for about 90% of C_l at $l = 10$, and about 50% at $l = 30$ [7].

The Sachs-Wolfe effect is determined by the curvature perturbation. In the case $\Omega_0 = 1$ it is given by [23, 46, 54]

$$\Delta T(\mathbf{e})/T = -\frac{1}{5}\mathcal{R}(\eta_0 \mathbf{e}) \tag{68}$$

where $\eta_0 = 2(a_0 H_0)^{-1}$ is the coordinate distance of the edge of the observable universe, taken to correspond to the surface of last scattering. (The right hand side is usual given as $\frac{1}{3}\Phi(\eta_0 \mathbf{e})$ where Φ is the peculiar gravitational potential.) Using Eq. (19) the multipoles are therefore given by

$$a_{lm} = -\frac{1}{5}\int_0^\infty dk \mathcal{R}_{klm}\sqrt{\frac{2}{\pi}}k j_l(\eta_0 k) \ . \tag{69}$$

Using Eq. (24), the mean square multipoles C_l are therefore given by

$$C_l = \frac{4\pi}{25}\int_0^\infty \frac{dk}{k} j_l^2(\eta_0 k)\, \mathcal{P}_\mathcal{R}(k) \tag{70}$$

For $\Omega_0 < 1$, keeping for the moment only sub-curvature modes, the Sachs-Wolfe effect is given by [4, 33, 34]

$$a_{lm} = -\int_0^\infty dq \mathcal{R}_{klm} q I_{kl} \tag{71}$$

$$q I_{kl} = \frac{1}{5}\Pi_{kl}(\eta_0) + \frac{6}{5}\int_0^{\eta_0} dr \Pi_{kl}(r)F'(\eta_0 - r) \tag{72}$$

Here $\mathcal{R}_{klm}$ is evaluated well before Ω breaks away from 1 (when it is constant), F is given by Eq. (61) and $\eta_0 = 2(1-\Omega_0)^{1/2}$ is again the coordinate distance of the edge of the observable universe. The mean square multipoles are given by

$$C_l = 2\pi^2 \int_1^\infty \frac{dk}{k}\mathcal{P}_\mathcal{R}(k) I_{kl}^2 \tag{73}$$

When $k \to 1$, I_{kl} tends to a finite and nonzero limit for each l. This means that the C_l are finite provided that $q^2\mathcal{P}_\mathcal{R} \to 0$.

Two things should be noted about the regime $q \to 0$ in the curved space case. First, all multipoles receive contributions from this regime; in contrast with the flat case the quadrupole does not dominate as is claimed in [41]). Second, the limit $q \to 0$ corresponds to scales of order the curvature, not to infinitely large scales as is claimed in [35, 41]. Because of this last fact, one cannot tolerate a divergence of the C_l as $q \to 0$ (unless the curvature scale is much bigger than any scale of interest in which case one is back to the flat-space case).

5.3 Inflation and horizon exit

It is widely supposed that the hot big bang is preceded by an era of inflation, during which gravity is by definition repulsive. A very attractive hypothesis is that the curvature perturbation originates as a vacuum fluctuation during inflation, so that the ensemble average appearing in the definition of the spectrum (Eq. (23)) is just the vacuum expectation value. Made originally for the case $\Omega_0 = 1$ [55, 51, 43, 56, 57], this hypothesis was later extended to the case $\Omega_0 < 1$ by Lyth and Stewart [4]. Before discussing it, let us see how inflation works with special reference to the case $\Omega_0 < 1$.

The Hubble distance H^{-1} is usually termed the horizon (to be distinguished from the particle horizon), and the comoving length scale a/k associated with a given mode is said

to be outside the horizon if $aH/k > 1$, and inside the horizon if $aH/k < 1$. The evolution of perturbations outside the horizon is very simple, because it is not affected by causal processes. Instead, the perturbation evolves independently in each comoving region [46].

Super-curvature scales, $a/k > 1$, are always outside the horizon (from Eq. (1)), but sub-curvature scales can be either outside or inside it. In the usual cosmology where gravity is attractive, $aH \equiv \dot{a}$ decreases with time and at each epoch some scale is entering the horizon. The Hubble scale H_0^{-1} is entering the horizon now, and smaller scales entered the horizon earlier. Also, from Eq. (1), Ω is driven away from 1 as time passes, so that $|1 - \Omega|$ must have been extraordinarily small at early times even if it is not small now.

Inflation may be defined as an early era of repulsive gravity, when $aH \equiv \dot{a}$ increases with time, and it is widely supposed that such an era preceded the hot big bang. At each epoch during inflation some scale is leaving the horizon, and as time goes by Ω is driven towards 1. The standard assumption is that inflation occurs because the scalar field potential dominates the energy density, which falls slowly with time owing to the evolution of one of the scalar fields, termed the inflaton field. Constant energy density corresponds to $\Omega \propto H^{-2}$, and combining this dependence with Eq. (1) gives, for the case $\Omega < 1$,

$$a = \hat{H} \sinh(\hat{H}t) \tag{74}$$

and

$$H = \hat{H} \coth(\hat{H}t) \tag{75}$$

After Ω has been driven to 1, H achieves the almost constant value $\hat{H}$, and

$$a \propto \exp(Ht) \tag{76}$$

It is related to the scalar field potential V (in turn practically equal to the energy density) by

$$\hat{H}^2 = \frac{8\pi}{3} m_{Pl}^{-2} V \tag{77}$$

(After many Hubble times the Hubble constant could vary appreciably, in which case $\hat{H}$ denotes the value before this happens.)

This evolution of the scale factor is modified if the energy density is rapidly decreasing with time; in particular, inflation might begin with a 'coasting epoch' during which Ω is almost constant [4, 58]. We shall not consider that case.

In the case $\Omega_0 = 1$, inflation is usually held to solve at least three problems that arise if the hot big bang extends back to the Planck scale. Let us briefly recall them, and consider whether they are still solved if $\Omega_0 < 1$. For the moment we are discounting the bubble nucleation model of inflation.

1. *The harmful relic problem* Without inflation it is difficult to avoid harmful relics of the early universe like monopoles or gravitinos. The possibility of avoiding them by inflation does not depend on the value of Ω_0.

2. *The flatness problem* As we go back through time during the hot big bang, Ω is driven very close to 1 so that we have a fine tuning problem, the flatness problem. It is solved provided that inflation lasts long enough that Ω has been driven away from 1 again as we go back to the beginning of inflation. From Eq. (74) (or its $\Omega > 1$ counterpart), the era at the beginning of inflation during which Ω is significantly different from 1 has a typical duration of order $\hat{H}^{-1}$ where $\hat{H}$ is related to the scalar field potential by Eq. (77). From Eq. (1), Ω has its present value Ω_0 at the epoch during inflation when the observable universe (or to be precise, the comoving scale presently equal to the Hubble distance) leaves the horizon. Thus, Ω_0 will be different from 1 if this epoch occurs near the beginning of inflation on the timescale $\hat{H}^{-1}$, but close to 1 otherwise.

3. *The homogeneity (horizon) problem* Without inflation, the observable universe is far outside the horizon (Hubble distance) at early times. This means that causal processes cannot determine the initial conditions, which is usually held to be a problem, termed the 'horizon problem'. If Ω_0 is close to 1, the observable universe is typically far inside the horizon at the beginning of inflation, which solves the horizon problem and is usually said to 'explain' the homogeneity of the observable universe. Inflation with $\Omega_0 < 1$ cannot solve the horizon problem because the observable universe (or to be precise the comoving length presently equal to the Hubble distance) never occupies less than a fraction $1 - \Omega_0$ of the Hubble distance.

However, no causal mechanism has ever been proposed for actually establishing homogeneity at the beginning of inflation, even after the horizon problem has been solved. It seems to us therefore that the 'horizon problem' is a red herring, and that one should therefore look elsewhere for an explanation of the homogeneity of the universe. For the case $\Omega_0 \simeq 1$ a fruitful avenue seems to be the following [46]. As smaller and smaller scales are considered one expects to find homogeneity below some minimum scale, but this is not the Hubble distance even though that is the only scale available at the classical level. Rather it is the scale, available only at the quantum level,

$$\rho^{-1/4} = \left(\frac{3}{8\pi}\right)^{1/4} \left(\frac{H}{m_{Pl}}\right)^{1/2} H^{-1} \tag{78}$$

(we are setting $\hbar$ as well as c equal to 1). Indeed, within a volume with this radius, even the vacuum fluctuation of a massless scalar field generates energy density *and pressure* of order $\rho^{1/4}$, which would spoil inflation. As in the case of flat spacetime this vacuum contribution to the energy density is to be discounted (ie., one has to solve the cosmological constant problem by fiat at our present level of understanding). But one cannot allow a significant occupation number for the particle states defined on this vacuum. In other words, if $\Omega_0 = 1$ the universe has to be *absolutely homogeneous* at the classical level, on scales smaller than $\rho^{-1/4}$. This guarantees the homogeneity of the observable universe at the classical level, provided that inflation starts at least $[\ln(m_{Pl}/H_1)]$ Hubble times before the observable universe leaves the horizon, where H_1 is the value of H at this latter epoch. In order to respect the isotropy of the cmb one requires $(H_1/m_{Pl})^{1/2} \lesssim 10^{-3}$ [59], and the bound is saturated in typical models of inflation. Thus, homogeneity of the observable universe is typically guaranteed if it leaves the horizon more than 7 or so Hubble times after the beginning of exponential inflation [46].

If $\Omega_0 < 1$ it is unclear how to define the vacuum as we discuss below, but with the mathematically simple conformal vacuum the vacuum fluctuation again generates an energy density and pressure of order d^{-4} on the scale d. The criterion that this should not spoil the inflationary behaviour Eq. (74) is that d^{-4} be much less than the critical density, which requires as before $d \gtrsim (H/m_{Pl})^{1/2} H^{-1}$. But now the observable universe is never far inside the Hubble distance H^{-1}, so its homogeneity is not guaranteed by this type of argument. A different avenue would be to invoke quantum cosmology, along the lines of [60] which however deals only with the case $\Omega_0 > 1$.

The bubble nucleation model of inflation

All of the above discussion assumes a classical evolution for the inflaton field, leading to a smooth evolution of $\Omega(t)$. It might happen, however, that the scalar field potential allows quantum tunneling in scalar field space at some point during inflation. In that case a bubble of scalar field can form, whose interior is an $\Omega \ll 1$ universe [25, 26, 27, 28]. Provided that the scalar field potential is still flat enough, Ω will again be driven to 1.

If Ω_0 turns out to be less than 1 the bubble nucleation model will be very attractive. Homogeneity is automatic. Also, Ω_0 is determined by the form of the scalar field potential and can easily be less than 1 [26]. Assuming the usual 'chaotic' scenario for the beginning of inflation [24, 61, 62], the inflaton field rolls slowly down a valley in scalar field space, and then the bubble nucleation model might correspond to sideways tunneling out of this valley [63]. The only problem would be to find a potential of the required form that looks sensible in the context of modern particle theory; as has recently been pointed out [58, 64], this constraint makes it difficult even to find a potential that leads to ordinary, non-tunneling inflation.

5.4 Sub-curvature contributions and the vacuum fluctuation

During inflation, the curvature perturbation is related to the perturbation $\delta\phi$ of the inflaton field by [57, 4, 46]

$$\mathcal{R} = -(H/\dot{\phi})\delta\phi \tag{79}$$

where the dot denotes differentiation with respect to time t. This expression holds at all epochs, not just when $\mathcal{R}$ is constant. In it, $\delta\phi$ is defined [57] on hypersurfaces which have zero perturbation in their curvature scalar (it is often called the 'gauge invariant' inflaton field perturbation).

A very attractive hypothesis is that $\delta\phi$ originates as a vacuum fluctuation, so that the ensemble average appearing in the definition of the spectrum (Eq. (23)) is just the vacuum expectation value [55, 51, 43, 56, 57, 65].

The vacuum fluctuation during inflation also generates a spectrum of gravitational waves, which is well understood for the case $\Omega_0 = 1$ [66], and is under investigation for the case $\Omega_0 < 1$ [67]. We will not consider it here.

To calculate the vacuum fluctuation one uses quantum field theory in negatively curved space [19, 68], and the first step in setting up this theory is to expand $\delta\phi$ in terms of the *sub-curvature* mode functions. In this context there is no question of including additional modes, because many results of quantum field theory (such as the vanishing of field commutators outside the light cone) depend essentially on the fact that one is using a complete orthonormal set. As a result the spectrum predicted by the vacuum fluctuation will include only sub-curvature modes.

The same restriction holds for the fluctuation in any quantum state that is homogeneous (with respect to the group of coordinate transformations leaving the distance between each pair of points invariant). But one can give the inflaton field perturbation any desired stochastic properties by choosing a suitable quantum state (pure or mixed), and in particular one can generate an arbitrary homogeneous Gaussian perturbation. The absence of super-curvature modes in the vacuum fluctuation prediction is not a feature of quantum field theory *per se*.

The coefficients in the mode expansion of the quantum field ϕ_{klm} satisfy the classical field equation (in the Heisenberg representation), which fixes them up to a one-parameter ambiguity once a convention is made for their normalisation. Breaking this ambiguity is equivalent to defining the vacuum. In the case $\Omega_0 = 1$ each mode starts out well inside the horizon, where the spacetime curvature is negligible. In that case the vacuum is defined to be the usual flat spacetime vacuum, and assuming the usual slow roll conditions one finds [43]

$$\mathcal{P}_{\mathcal{R}}(k)^{1/2} = \frac{8}{m_{Pl}^3}\sqrt{\frac{2\pi}{3}}\frac{V^{3/2}}{V'} \tag{80}$$

In this expression $V(\phi)$ is the inflaton potential, and the right hand side is to be evaluated at the epoch of horizon exit $k = aH$. It gives an almost scale-independent result for typical models of inflation.

In the case $\Omega_0 < 1$, without bubble nucleation, it is not clear how to define the vacuum because a given scale is never far inside the horizon. The mathematically simplest choice is the 'conformal vacuum', and using it one finds (after suitably generalising the slow roll conditions) that $\mathcal{P}_\mathcal{R}$ is still given by the above expression [4, 52]. (For the special case of a linear potential this result has been reproduced recently, using a different calculational technique [39].[6])

In the bubble nucleation model the quantum state of the inflaton field perturbation inside the bubble can be calculated [27, 26, 67], and it is found *not* to be in the conformal vacuum. As a result [26, 28], $\mathcal{P}_\mathcal{R}(k)$ is multiplied by a factor $\coth(\pi q)$ compared with Eq. (80).

Comparison of the vacuum fluctuation with observation

If $\Omega_0 = 1$, Eq. (70) with a scale independent $\mathcal{P}_\mathcal{R}$ gives

$$l(l+1)C_l = \frac{2\pi}{25}\mathcal{P}_\mathcal{R} \tag{81}$$

The prediction is that the left hand side is scale independent, which is consistent with observation [6, 7]. There is however room for considerable scale dependence; defining n by $\mathcal{P}_\mathcal{R} \propto q^{n-1}$, the allowed range is $0.6 \lesssim n \lesssim 1.4$. The magnitude is small,

$$l(l+1)C_l = 8.05 \times 10^{-10}\frac{Q_{\text{RMS-ps}}}{20\,\mu\text{K}} \simeq 8.0 \times 10^{-10} \tag{82}$$

which corresponds to

$$\mathcal{P}_\mathcal{R} \simeq 3 \times 10^{-9} \tag{83}$$

For $\Omega_0 < 1$ Eq. (73) has to be evaluated numerically. With the flat spectrum coming from the conformal vacuum assumption, $l(l+1)C_l$ has in general a negative slope for $0.1 < \Omega_0 < 1$ (for l in the range $l \lesssim 10$ where the Sachs-Wolfe effect dominates, and in which curvature can be significant) [4, 39, 37, 69]. With present data the slope is not strong enough to rule out any value of Ω_0, though better data will probably rule out low values.

As already noted, the spectrum is not flat in the bubble nucleation model, but rather is proportional to $\coth(\pi q)$. It turns out however [28] that if Ω_0 is substantially below 1 the integral in Eq. (72) dominates, with I_{kl}^2 peaking at $q^2 \gtrsim 1$ even for the quadrupole (and at higher values for higher multipoles). As a result the bubble nucleation prediction is not significantly different from the flat spectrum prediction, when cosmic variance is taken into account.

Power law parameterization of the density perturbation spectrum

The spectrum $\mathcal{P}_\mathcal{R}$ of the curvature perturbation is directly related to the Sachs-Wolfe effect, and from this viewpoint the assumption that it is effectively flat seems natural. This assumption was not, however, the assumption made in the literature before the (very recent) advent of the vacuum fluctuation prediction. Rather, it was assumed that the spectrum P_δ (defined by Eq. (33)) of the density perturbation is proportional to q. This choice is equivalent to the flatness of $\mathcal{P}_\mathcal{R}$ for $\Omega_0 = 1$, but otherwise it is equivalent to

$$\mathcal{P}_\mathcal{R} \propto \frac{q^2(1+q^2)}{(4+q^2)^2} \tag{84}$$

---

[6]The authors of [39] do not establish the identity of their result with the earlier one, but it follows by evaluating Eq. (2) of [37] (multiplied by $16\pi/m_{Pl}^2$ to bring the conventions of [39] into line with the usual ones) during matter domination before Ω breaks away from unity. To do this one has to replace $1 - \Omega_0$ by $1 - \Omega \ll 1$ in the quantity W_1/c_1, leading to [23] $W_1/c_1 \to (2/5)(aH)^{-2}$. Remembering that the energy density scales like a^{-3} during matter domination and like a^{-4} during radiation domination, one then indeed reproduces the spectrum of the energy density given by Eqs. (63) and (80).

The right hand side tends to 1 in the limit $q^2 \to \infty$ of negligible curvature, but is much less than 1 for $q^2 \lesssim 4$. With this parameterization, $l(l+1)C_l$ acquires a *positive* slope [36] for $0.1 < \Omega_0 < 1$, comparable in magnitude with the negative slope of the vacuum fluctuation prediction.

The main cause of this difference is the factor $(4+q^2)^2$ coming from the relation between the density and curvature perturbations, and a similar result would probably be obtained if $\mathcal{P}_\delta$ were used instead of P_δ, or k instead of q. In other words, the predicted C_l can be regarded as coming simply from a *linearly rising density perturbation spectrum*, as opposed to a *flat curvature perturbation spectrum*. These two parameterizations are equivalent if Ω_0 is equal to 1, but if Ω_0 is significantly smaller than 1 the first one gives less power on small scales, leading to a significantly different prediction for the C_l's. The central values of the present data points lie between the two predictions, and the present error bars are big enough that they are indistinguishable. But in the future the data should be able to distinguish between the two parameterizations, ruling out one or both of them for small values of Ω_0.

5.5 Super-curvature scales and the Grishchuk-Zeldovich effect

Like any statement in physics, the statement that the inflaton field is in the vacuum will be at best approximate, and its validity will presumably depend on the scale under consideration. When considering departures from it there is no reason to exclude super-curvature scales.

The contribution of super-curvature scales to the mean square multipole C_l is just the extension of Eq. (73) to super-curvature scales,

$$C_l^{\mathrm{SC}} = 2\pi^2 \int_0^1 \frac{\mathrm{d}k}{k} \mathcal{P}_\mathcal{R}(k) I_{kl}^2 \tag{85}$$

Requiring that the super-curvature contribution be no bigger than the total gives (Eq. (82))

$$l^2 C_l^{\mathrm{SC}} \lesssim 10^{-10} \tag{86}$$

The quantities I_{kl}^2 have not been calculated in the super-curvature regime $0 < k^2 < 1$, but as discussed below they are proportional to k^2 near $k = 0$, and we noted earlier that in the regime $k^2 > 1$ they all peak at a value $k^2 \gtrsim 2$. It therefore seems reasonably to suppose that for a fairly flat spectrum $\mathcal{P}_\mathcal{R}(k)$, the super-curvature regime does not contribute much to the mean square multipoles C_l. In that case Eq. (86) will provide no significant constraint on a flat spectrum. In other words, it will be difficult to detect the cutoff below $k^2 = 1$ that the vacuum fluctuation predicts. (The even more difficult task of finding an observational signal for the this cutoff *without* assuming that the spectrum is flat is discussed in [70].)

Now suppose, in contrast, that the spectrum rises sharply on some very large scale. Then there might be a big curvature perturbation, which would however have a big correlation length and so have a very small spatial gradient. If the correlation length is big enough, the gradient will be small enough to ensure that the curvature perturbation has no significant effect on the cmb anisotropy, even if it is quite is large. How big does the correlation length have to be for this to happen?

For $\Omega_0 = 1$ this question was asked and essentially answered by Grishchuk and Zeldovich [10]. We will now briefly recall their argument, using the precise concept of the spectrum of the curvature perturbation in place of their more qualitative discussion. Then we will generalize it to the case $\Omega_0 < 1$, and finally discuss its physical significance in both cases.

5.6 The Grishchuk-Zeldovich effect ($\Omega_0 = 1$)

From Eq. (31), a flat spectrum $\mathcal{P}_\mathcal{R}$ gives a logarithmically divergent result for $\langle \mathcal{R}^2 \rangle$, but since $\mathcal{P}_\mathcal{R} \sim 10^{-9}$ one has to go to a huge scale to see any effect. Taking the comoving

small-scale cutoff to be the Hubble distance at the end of inflation (which is equivalent to the usual procedure of dropping the contribution of the vacuum fluctuation to the energy density in flat spacetime), this scale as a multiple of the Hubble distance is

$$\frac{a_0 H_0}{k_{\mathrm{HU}}} \sim \exp(10^9 - 60) \sim \exp(10^7) \tag{87}$$

If as expected the spectrum increases somewhat with scale this estimate will be sharply reduced, but it will still be a big number in typical models of inflation.

A possible divergence of the geometry distortion associated with a nearly flat spectrum is interesting, because it suggests that the universe might be fractal on very large scales [61]. It cannot however be explored observationally, because if the spectrum is fairly flat, large scales give a negligible contribution to the cmb anisotropy. Here we are concerned with the quite different possibility, that the spectrum might rise sharply on some not-so-huge but still very large scale q_{VL}. We therefore suppose that the spectrum has the form

$$\mathcal{P}_{\mathcal{R}}^{\mathrm{VL}} \simeq \delta(\ln k - \ln k_{\mathrm{VL}})\langle \mathcal{R}^2 \rangle \tag{88}$$

Such a contribution might originate from the vacuum fluctuation if the inflaton potential has a suitable form, but more plausibly it would arise because the vacuum assumption failed, or in other words because there were inflaton particles with momentum $k \lesssim k_{\mathrm{VL}}$.

Before calculating its effect, let us spell out the physical significance of this contribution. We will make the natural assumption that the homogeneous Gaussian perturbation under consideration exists in a patch around us, whose size is much bigger than the correlation length

$$d_{\mathrm{VL}} = a_0/k_{\mathrm{VL}} \tag{89}$$

The curvature perturbation is more or less constant in a region of size d_{VL}. According to Gauss-Bonnet theorem, the distortion in geometry in this region (measured for instance by the departure of the sum of the angles in a geodesic triangle spanning it from 2π) is of order its cross-sectional area times the perturbation in the curvature scalar. From Eq. (31) this is of order

$$d_{\mathrm{VL}}^2 \delta R^{(3)} \sim \langle \mathcal{R}^2 \rangle^{1/2} \tag{90}$$

In some patches the curvature perturbations will be positive, and in others negative. It does not make sense to consider a value $\langle \mathcal{R}^2 \rangle$ bigger than 1, because then the regions of positive curvature would close on themselves.

The geometry distortion in the observable universe is of order $a_0^2 \delta R^{(3)}$, and so smaller by a factor k_{VL}^2. As the correlation length increases with fixed $\langle \mathcal{R}^2 \rangle$, the geometry distortion in the observable universe decreases and so does the spatial gradient For both these reasons, the anisotropy caused by a given value of $\mathcal{P}_{\mathcal{R}}$ decreases. Let us calculate it.

Since $j_l(x) \sim x^l$ for small x, one sees from Eq. (70) that the quadrupole dominates. Using $j_2(x) = x^2/15$ one finds

$$6C_2^{\mathrm{VL}} = \frac{4\pi}{25} \frac{16}{15^2} \left(\frac{k_{\mathrm{VL}}}{a_0 H_0}\right)^4 \langle \mathcal{R}^2 \rangle \tag{91}$$

The quadrupole measured by COBE is not significantly in excess of the typical values $l(l+1)C_l \simeq 8 \times 10^{-10}$ of the other multipoles (in fact it is somewhat smaller), so we conclude that C_2^{VL} is absent at this level. This means that

$$d_{\mathrm{VL}} > 70\langle \mathcal{R}^2 \rangle^{1/4} H_0^{-1} \tag{92}$$

As Grishchuk and Zeldovich pointed out, this bound on $\langle \mathcal{R}^2 \rangle$ becomes weaker as the scale increases. It can be of order 1 provided that (cf. [71])

$$\frac{a_0 H_0}{k_{\mathrm{VL}}} > 70 \tag{93}$$

A bound similar to this one was already implied by upper limits on the quadrupole that existed two decades ago. But before COBE measured the actual values of the low multipoles there was always the possibility that the Grishchuk-Zeldovich effect might be present (ie., that the quadrupole might stick out above the other multipoles), indicating a big curvature perturbation on some very large scale. The somewhat disappointing fact that the Grishchuk-Zeldovich is absent seems not to have been noted anywhere in the copious literature on the COBE observations.

5.7 The super-curvature scale Grishchuk-Zeldovich effect

To generalize the Grishchuk-Zeldovich effect to $\Omega_0 < 1$ one needs to take spatial curvature into account, and to note that the limit of large scales corresponds to $k \to 0$, not $q \to 0$. This has not been done to date. The only relevant publications of which we are aware are [71] where spatial curvature is ignored, and [35, 41] where the appropriate limit is incorrectly assumed to be $q \to 0$. In contrast with the case $\Omega_0 = 1$, the Grishchuk-Zeldovich effect for the case $\Omega_0 < 1$ could come only from a non-vacuum contribution.

Consider therefore a contribution on very large super-curvature scales, and represent it by

$$\mathcal{P}_{\mathcal{R}}^{\mathrm{VL}} \simeq \delta(\ln k - \ln k_{\mathrm{VL}})\langle \mathcal{R}^2 \rangle \tag{94}$$

We need to understand the physical significance of the corresponding curvature perturbation, which is different from the case $\Omega_0 = 1$. From Section 4.1, the correlation length is now

$$d_{\mathrm{VL}} = a_0 / k_{\mathrm{VL}}^2 \tag{95}$$

The curvature perturbation is practically constant in a region of this size. The *perturbation* in the geometry distortion of such a region is now

$$d_{\mathrm{VL}}^2 \delta R^{(3)} \sim k_{\mathrm{VL}}^{-4} \langle \mathcal{R}^2 \rangle^{1/2} \tag{96}$$

On the other hand there is now a distortion even in the absence of a perturbation, given by Eq. (2),

$$d_{\mathrm{VL}}^2 R^{(3)} \sim k_{\mathrm{VL}}^{-4} \tag{97}$$

Thus $\langle \mathcal{R}^2 \rangle^{1/2}$ now measures the fractional perturbation in the geometry distortion, not the distortion itself. However the requirement that regions of space should not close on themselves is still equivalent to $\langle \mathcal{R}^2 \rangle^{1/2} \lesssim 1$.

An equivalent way of viewing $\langle \mathcal{R}^2 \rangle^{1/2}$ is that it measures the geometry distortion of a region with size a_0. Since Ω_0 is not extremely small this is the same as saying that it measures the geometry distortion of the observable universe. As the correlation length d_{VL} is increased with $\langle \mathcal{R}^2 \rangle$ constant, the geometry distortion of the observable universe does not decrease as it does in the $\Omega_0 = 1$ case. One still expects, though, that for a given value of $\mathcal{P}_{\mathcal{R}}$ the effect on the cmb anisotropy will become smaller, because the spatial gradients become smaller. Let us see how to calculate it.

As $k \to 0$, $\Pi_{k0} \to 1$, but the other radial functions are proportional to k. The normalisation factor becomes

$$kN_{k1} \quad \to \quad N_1 \equiv \sqrt{\frac{2}{\pi}} \tag{98}$$

$$kN_{kl} \quad \to \quad N_l \equiv \sqrt{\frac{2}{\pi}} [\prod_{n=2}^{l}(n^2 - 1)]^{-1/2} \quad (l \geq 2) \tag{99}$$

and

$$\tilde{\Pi}_{k1}(r) \to \frac{k^2}{4}\frac{1}{\sinh^2 r}[\sinh(2r) - 2r] \tag{100}$$

The other radial functions follow from the recurrence relation Eq. (15). It is convenient to define
$$\tilde{\Pi}_l \equiv \lim_{k \to 0} \tilde{\Pi}_{kl}/k \tag{101}$$

Using these results, the contribution to the mean square multipoles becomes
$$C_l^{\mathrm{VL}} = N_l^2 B_l^2 k_{\mathrm{VL}}^2 \langle \mathcal{R}^2 \rangle \tag{102}$$

where
$$B_l \equiv \frac{1}{5} \tilde{\Pi}_l(\eta_0) + \frac{6}{5} \int_0^{\eta_0} \tilde{\Pi}_l(r) F'(\eta_0 - r) \mathrm{d}r \tag{103}$$

When one increases the value of l under consideration, the scale above which these limits hold presumably becomes successively larger, so to actually calculate C_l^{VL} for a given value of k_{VL} one ought to use the full expression Eq. (85), but hopefully Eq. (102) will provide a reasonable estimate for small l. Since N_l and B_l are roughly of order 1 for low multipoles, it says very roughly that
$$C_l^{\mathrm{VL}} \sim k_{\mathrm{VL}}^2 \langle \mathcal{R}^2 \rangle \tag{104}$$

The absence of C_l^{VL} at the level 10^{-10} therefore implies very roughly
$$k_{\mathrm{VL}}^{-2} \gtrsim 10^{10} \langle \mathcal{R}^2 \rangle \tag{105}$$

Since $(a_0 H)^2 = 1 - \Omega_0$ is supposed not to be tiny, this result is roughly
$$d_{\mathrm{VL}} \gtrsim 10^{10} \langle \mathcal{R}^2 \rangle H_0^{-1} \tag{106}$$

In words, the conclusion is that if the fractional geometry distortion is of order 1, its correlation length must be more than 10^{10} Hubble distances. This result is not directly comparable with the $\Omega_0 = 1$ result, because it concerns the fractional, not the absolute, geometry distortion on the scale d_{VL}. The quantity that measures the absolute geometry distortion is $\tilde{\mathcal{R}}_{klm} \equiv k_{\mathrm{VL}}^4 \mathcal{R}_{klm}$, and in terms of this quantity
$$d_{\mathrm{VL}} \gtrsim 100 \langle \tilde{\mathcal{R}}^2 \rangle^{1/5} H_0^{-1} \tag{107}$$

We see that if the absolute geometry distortion, in a region whose diameter is equal to the correlation length, is of order 1, then the correlation length must be at least two orders of magnitude bigger than the Hubble distance. This is essentially the same as the $\Omega_0 = 1$ result.

These estimates have been derived from the fact that the Grishchuk-Zeldovich effect cannot be bigger than the observed values of the multipoles. As we have not actually calculated the l dependence of C_l^{VL} we cannot say that the effect is definitely absent as in the $\Omega_0 = 1$ case, because it might turn out that the dependence of C_{VL} mimics the dependence that of the data (roughly $C_l \propto l^{-2}$). It would be desirable to calculate the shape of C_l^{VL}, both to check that this does not happen and to check the assumption that B_l is of order 1.

In contrast with the case $\Omega_0 = 1$, the Grishchuk-Zeldovich effect is present in all of the low multipoles if $\Omega_0 < 1$. Indeed, it could even occur in multipoles $l \gtrsim 10$ in which case the Sachs-Wolfe approximation would become inadequate to investigate it and a full calculation would be necessary. The necessary formalism to perform such a calculation is already in place [34], and it has already been used for the sub-curvature modes [36, 37, 69]. The extension to the super-curvature modes raises no new issue of principle.

5.8 The physical significance of the Grishchuk-Zeldovich effect

In some of the literature [71, 41], the absence of the Grishchuk-Zeldovich effect at a given level has been regarded as evidence that the smooth patch of the universe which we occupy extends beyond the edge of the observable universe. As we now explain, this is not the case.

To make the simplest point first, it is clear from Eqs. (71) and (72) that the multipoles a_{lm} of the cmb anisotropy depends only on the curvature perturbation within the observable universe. This remains true when we consider their ensemble mean squares C_l. Strictly interpreted, the Grishchuk-Zeldovich effect just explores the effect of very small spatial gradients of the curvature perturbation, within the observable universe, on the hypothesis that the curvature perturbation is a typical realization of a homogeneous Gaussian random field. Recall that in this context 'homogeneous' means that the correlation function of the curvature perturbation depends only on the distance between two points, not on their location; but we are still talking about locations within the observable universe.

If this hypothesis is indeed correct within the observable universe, one expects it to remain correct in some larger region. *If* this region is sufficiently big, one can introduce the concept of a correlation length as we did in the above discussion. By definition, the correlation function is more or less constant out to a distance of order the correlation length, only then falling off. Clearly 'sufficiently big' means bigger than the correlation length. But we will never know whether this picture is correct, because we will never know what lies beyond the edge of the observable universe (except by waiting for it to gradually recede, in comoving distance units).

Homogeneity of the perturbation corresponds to the spectrum (defined by Eq. (47)) being independent of l and m. One can reasonably expect this property to fail when k becomes so small that the corresponding distance a_0/k_{VL}^2 becomes bigger than the size of the smooth patch of the universe around us, within which the perturbation is homogeneous. In that case the *presence* of the Grishchuk-Zeldovich effect on a given scale would suggest, though not really prove, that the smooth patch extends to the corresponding distance. But its absence says nothing.

Beyond the smooth patch might be regions of the universe where the 'perturbations' become so big that it makes no sense to talk about a Robertson-Walker universe. If so the patch discussed in the last paragraph will have a periphery, within which the typical magnitude of the perturbations becomes bigger as one moves outwards (in contrast with the region within the patch, where the typical magnitude is by definition the same everywhere). The Grishchuk-Zeldovich effect tells us absolutely nothing about this periphery.

For the case $\Omega_0 = 1$ one might argue that information about the periphery is available, by taking the density perturbation to be the primary quantity rather than the curvature perturbation. From this viewpoint the large density perturbation in the periphery will generate a large curvature perturbation in the observable universe, through the usual 'Coulomb law' solution of the Poisson equation Eq. (64), unless there is an accidental cancellation. However there does not seem to be any justification for it, and it does not work for $\Omega_0 < 1$ because according to Eq. (63) the Poisson equation does not hold. Rather, the density perturbation and the curvature perturbation become essentially the same on scales much bigger than the curvature scale.

Inflation and the Grishchuk-Zeldovich effect

A separate issue is whether the absence of the Grishchuk-Zeldovich effect tells us anything about inflation. This is clearly the case only if the scale k_{VL} can be related to inflation. Reference [71], which deals with the case $\Omega_0 = 1$ (or at any rate ignores spatial curvature), accepts the usual dogma that 'the universe is smooth on some scale of order the Hubble distance at the beginning inflation'. Interpreting this to mean that $k_{VL} \sim aH$ at the

beginning of exponential ($\Omega = 1$) inflation, the absence of the Grishchuk-Zeldovich effect indeed tells us that inflation starts several Hubble times before the observable universe leaves the horizon. However, as discussed in Section 5.3 the usual dogma does not have any clear justification.

6 Conclusion

In this article we have drawn the attention of physicists to the incompleteness of the standard mode expansion for cosmological perturbations in an $\Omega_0 < 1$ universe. In order to generate the most general homogeneous Gaussian random field one should use Eq. (47), which runs over all negative eigenvalues $-k^2$, whereas the standard expansion keeps only the modes with $k^2 > 1$. We have called these sub-curvature modes, because they vary appreciably over a distance less than the curvature scale, and we have called the modes with $0 < k^2 < 1$ super-curvature modes.

The fact that super-curvature modes are needed to generate the most general perturbation has been known to mathematicians for about half a century, so that their omission by cosmologists constitutes a remarkable failure of communication between the worlds of mathematics and science. This omission leads to perturbations which are practically uncorrelated on scales bigger than the curvature scale. In contrast, a mode with $k^2 \ll 1$ corresponds to a perturbation with correlation length k^{-2} in units of the curvature scale.

What nature has chosen to do is of course another question. For the case $\Omega_0 = 1$ the standard assumption is that the perturbations originate as the vacuum fluctuation of the inflaton field, and in 1990 this assumption was extended to the case $\Omega_0 < 1$ [4]. The mode expansion of a quantum field runs only over sub-curvature modes, since they form a complete orthonormal set for square integrable function. As a result the vacuum fluctuation generates a perturbation which includes only these modes.

Of course this is not a feature of quantum field theory *per se*, but of the assumption that the inflaton field is in the vacuum. This assumption might break down below some small value of k, corresponding to a correlation length much bigger than the curvature scale. We therefore ask whether a big perturbation with a very large correlation length could be detected through the cmb anisotropy. For the case $\Omega_0 = 1$, this question was asked and answered in 1978 by Grishchuk and Zeldovich [10], and here we have extended their discussion to the case $\Omega_0 < 1$ by including the super-curvature modes. We have given a formula for the cmb anisotropy due to a mode with $k \ll 1$, and have estimated its magnitude. By requiring that it be no bigger than the observed anisotropy we have estimated a lower limit on the correlation length for a perturbation of given magnitude. As in the case $\Omega_0 = 1$, the correlation length must be more than about two orders of magnitude bigger than the size of the observable universe, if the geometry distortion is of order 1 in a region whose size is equal to the correlation length.

In contrast with the case $\Omega_0 = 1$, the Grishchuk-Zeldovich effect is present in all multipoles, not just in the quadrupole. It would be interesting to evaluate its l dependence, if only to check that it does not mimic the observed dependence which is usually interpreted as coming entirely from the vacuum fluctuation.

Appendix

This appendix gives some mathematical results, in the sort of language that we as physicists are accustomed to. It deals with both the spherical expansion used in the text, and with an expansion using coordinates that slice space into flat surfaces which is more like the flat-space Fourier series. At the risk of being pedantic we give a rather full treatment, because

even in the flat-space case there does not seem to be a reference that explains the basic concepts in a way that is accessible to most physicists.

We refer the reader seeking a rigorous but more abstract treatment to [11].

The Fourier expansion

In flat space the simplest approach is to use the Fourier expansion. In comoving coordinates $\mathbf{r} \equiv (x, y, z)$ it is

$$f(\mathbf{r}) = (2\pi)^{-3/2} \int \mathrm{d}^3\mathbf{q} f_{\mathbf{q}} \exp(i\mathbf{q}.\mathbf{r}) \tag{108}$$

The orthonormality relation is

$$(2\pi)^{-3} \int \mathrm{d}\mathcal{V} \exp(-i\mathbf{q}.\mathbf{r}) \exp(i\mathbf{q}'.\mathbf{r}) = \delta^3(\mathbf{q}' - \mathbf{q}) \tag{109}$$

where $\mathrm{d}\mathcal{V} = \mathrm{d}^3\mathbf{r}$ is the volume element.

A Gaussian perturbation is obtained by assigning independent Gaussian probability distributions to the real and imaginary parts of the coefficients $f_{\mathbf{q}}$, but we need to ensure translation and rotation invariance of the correlation function. Translation invariance is equivalent to the real and imaginary parts having the same distributions (same mean squares), for the following reason. Because $\exp[i(\mathbf{q}.\mathbf{x})] \exp[i(\mathbf{q}'.\mathbf{x}')]$ is a function of $\mathbf{x} - \mathbf{x}'$ only if $\mathbf{q} = -\mathbf{q}'$, translation invariance is equivalent to introducing a correlation between $f_{\mathbf{q}}$ and $f_{-\mathbf{q}}$ only, which because of the reality condition $f_{\mathbf{q}}^* = f_{-\mathbf{q}}$ means a correlation between $f_{\mathbf{q}}$ and $f_{\mathbf{q}}^*$ only. This means that the phases of $f_{\mathbf{q}}$ must be uncorrelated, which indeed means that the real and imaginary parts of $f_{\mathbf{q}}$ must have the same distribution.[7] Let us therefore define the spectrum by

$$\langle f_{\mathbf{q}}^* f_{\mathbf{q}'} \rangle = \frac{2\pi^2}{q^3} \mathcal{P}_f(q) \delta(\mathbf{q} - \mathbf{q}') \tag{110}$$

The correlation functions is then

$$\xi_f(r) = (2\pi)^{-3} 2\pi^2 \int \frac{\mathrm{d}^3\mathbf{q}}{q^3} \exp(-i\mathbf{q}.\mathbf{r}) \mathcal{P}_f(q) \tag{111}$$

Rotational invariance is clearly equivalent to the spectrum depending only on the magnitude of $\mathbf{q}$, not its direction. Performing the angular integration one obtains Eq. (29).

Using the well known expansion of a plane wave into spherical waves one can prove that the above definition of the spectrum is equivalent to the definition Eq. (24) in terms of the spherical expansion. The equivalence, and in particular the fact that the two definitions are the same except for the different delta functions, does not depend on the detailed form of the transformation between the spherical expansion and the Fourier expansion, but rather on the fact that it is unitary.

The spherical expansion

We first justify the claim made in the text, that the correlation function depends only on the distance between the points if the spectrum defined by Eq. (23) is independent of l and m. Let r, θ, ϕ and r', θ', ϕ' be the coordinates of a given point with respect to two different spherical coordinate systems. We saw earlier that the most general eigenfunction with eigenvalue $-(k/a)^2$ is a linear combination of the functions $Z_{klm}(r, \theta, \phi)$. This is of course true in any coordinate system. Since $Z_{klm}(r, \theta, \phi)$ and $Z_{klm}(r', \theta', \phi')$ are both eigenfunctions

---

[7]A direct way of seeing this is to work with the real form of the Fourier integral and note the identity $\cos a \cos b + \sin a \sin b = \cos(a - b)$

it follows that either of them can be expanded in terms of the other. Thus, there is a linear combination of the form

$$Z_{klm}(r,\theta,\phi) = \sum_{l'm'} U^k_{lml'm'} Z_{kl'm'}(r',\theta',\phi') \tag{112}$$

If the spectrum defined by Eqs. (23) and (43) is independent of l and m, the correlation function given by Eq. (25) or its super-curvature analogue will be invariant under transformations of the above form provided that the transformation matrix satisfies the unitarity property

$$\sum_{l''m''} U^k_{lml''m''}(U^k_{l'm'l''m''})^* = \delta_{ll'}\delta_{mm'} \tag{113}$$

Since the transformation takes a pair of points into arbitrary positions subject to the constraint that the distance between them is fixed, the correlation function will then depend only on this distance.

For the sub-curvature modes, unitarity follows from the fact that the transformation takes one orthonormal basis (for the subspace of eigenfunctions with a given eigenvalue) into another. Let us see this explicitly. Orthonormality for the whole L^2 space gives the the coefficients of the expansion as

$$\delta(q'-q)U^k_{lml'm'} = \int Z^*_{k'l'm'}(r',\theta',\phi')Z_{klm}(r,\theta,\phi)\mathrm{d}\mathcal{V} \tag{114}$$

where the primed coordinates are regarded as functions of the unprimed ones and the volume element is defined by Eq. (11). Now consider the inverse transformation,

$$Z_{kl'm'}(r',\theta',\phi') = \sum_{lm} V^k_{l'm'lm} Z_{klm}(r,\theta,\phi) \tag{115}$$

The coefficients are given by

$$\delta(q'-q)V^k_{l'm'lm} = \int Z^*_{k'lm}(r,\theta,\phi)Z_{kl'm'}(r',\theta',\phi')\mathrm{d}\mathcal{V}' \tag{116}$$

where now the unprimed coordinates are regarded as functions of the primed ones. But as the integration goes over all space one can just as well integrate over the unprimed coordinates and regarded the primed coordinates as the dependent ones. By comparing Eqs. (114) and (116) it follows that the transformation is indeed unitary,

$$\left(U^k_{lml'm'}\right)^* = V^k_{l'm'lm} \tag{117}$$

This proof of unitarity does not work for the super-curvature regime, because we invoked orthonormality to obtain Eq. (114) for the matrix element $U^k_{lml'm'}$. There is however an alternative expression that remains well behaved in the super-curvature regime, obtained by substituting into Eq. (112) the definition of the Z's, and remembering that the spherical harmonics Y_{lm} are a complete orthonormal set on the sphere. Choosing *any* sphere $r' =$constant this expression is

$$\Pi_{kl'}(r')U^k_{lml'm'} = \int \Pi_{kl}(r)Y_{lm}(\theta,\phi)Y^*_{l'm'}(\theta',\phi')\sin\theta'\mathrm{d}\theta'\mathrm{d}\phi' \tag{118}$$

where the unprimed coordinates are regarded as functions of the primed ones. (For sub-horizon modes the original expression Eq. (114) is recovered if we multiply both sides of Eq. (118) by $\Pi_{k'l'}(r')\sinh^2 r'\mathrm{d}r'$ and integrate over $0 < r' < \infty$, but continuing to imaginary q and q' causes the integrals on both sides to diverge at $r' = \infty$.) Similarly, choosing any sphere $r =$constant one has

$$\Pi_{kl}(r)V^k_{l'm'lm} = \int \Pi_{kl'}(r')Y_{l'm'}(\theta',\phi')Y^*_{lm}(\theta\phi)\sin\theta\mathrm{d}\theta\mathrm{d}\phi \tag{119}$$

In contrast with the original expressions Eqs. (114) and (116), the radial coordinate on the right hand side of these new expressions runs only over a finite range. When we analytically continue to imaginary q, the radial functions pick up a factor i which cancels, so the equality $\left(U^k_{lml'm'}\right)^* = V^k_{l'm'lm}$ that we established for real q remains valid. Note that this does not work for complex eigenvalues, indicating that they are not allowed.

The above discussion is a generalisation of the familar demonstration that a rotation around the origin acts on Y_{lm} with a *finite dimensional* unitary matrix acting on the m index alone. Because of the finite dimensionality, this provides a *rigorous* proof that invariance under such rotations is equivalent to the independence of the spectrum on m. The extension to arbitrary rotations and translations involves an infinite sum over l, and as we have not discussed its convergence the discussion is not rigorous. Sasaki and Tanaka [72] have recently demonstrated that the infinite sum over l that occurs in Eq. (25) is uniformly convergent, which makes the above derivation rigorous.

The flat-surface expansion

The flat-surface expansion uses coordinates defined by the line element [8, 17, 14]

$$\mathrm{d}l^2 = (a/z)^2(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2) \tag{120}$$

The surfaces of constant z are flat. Any point in space can be chosen as the point $x = y = 0$, $z = 1$ and curvature is negligible in the region around that point $|x| \ll 1$, $|y| \ll 1$, $|z-1| \ll 1$.

Since a sphere of infinite radius is flat, one can think of the flat surfaces $z = $constant as spherical wave fronts, originating from a point at $z = 0$. Note that these surfaces have an 'inside' and an 'outside' even in the limit where they are flat, because geodesic surfaces such as the 'equatorial plane' $\phi = \pi/2$ are *not* flat.

The virtue of these coordinates is that the form of the line element is invariant under the following transformation

$$\begin{aligned} x &\rightarrow C(x + X) \\ y &\rightarrow C(y + Y) \\ z &\rightarrow Cz \end{aligned} \tag{121}$$

With a suitable choice of the constants X, Y and C we can place one of the two points to which the correlation function refers at an arbitrary position, while leaving unchanged the form of the line element and therefore the geodesic distance to the other point.

If the point x', y', z' corresponds to the point $r = 0$ in the spherical coordinate system, one can orientate the axes so that [14]

$$\begin{aligned} x - x' &= z' \cos\phi \sinh r / \cosh(\eta - r) \\ y - y' &= z' \sin\phi \sinh r / \cosh(\eta - r) \\ z &= z' \cosh\eta / \cosh(\eta - r) \end{aligned} \tag{122}$$

where $\tanh\eta \equiv \cos\theta$.

Now consider the mode expansion. We look for eigenfunctions of the form

$$W_{k\mathbf{q}_\perp} = \int \mathrm{d}\mathbf{q}_\perp^2 F_{k\mathbf{q}_\perp}(z) e^{i\mathbf{q}_\perp \cdot \mathbf{x}} \tag{123}$$

where $\mathbf{x}$ is the vector with components x, y. Substituting this expression into the Laplacian gives a second order equation for $F_{k\mathbf{q}_\perp}$. One of its two linearly independent solutions is [17, 14] $z K_{iq}(q_\perp z)$ where K is the modified Bessel function. This solution vanishes at $z = \infty$, and up to normalisation is the only solution with that property. It also vanishes at

$z = 0$ (spatial infinity in opposite direction) for real k. Note that K_{iq} is real for real q^2 and imaginary for imaginary q.

We will define the sub-curvature modes by

$$W_{k\mathbf{q}_\perp}(\mathbf{x}, z) \equiv N(k)e^{i\mathbf{q}_\perp \cdot \mathbf{x}}zK_{iq}(q_\perp z) \tag{124}$$

where

$$N^2(k) = \frac{q\sinh(\pi q)}{2\pi^4} \tag{125}$$

The normalisation has been chosen to satisfy the orthonormality condition

$$\int W_{k\mathbf{q}_\perp}(\mathbf{x}, z)W_{k'\mathbf{q}'_\perp}(\mathbf{x}, z)\mathrm{d}\mathcal{V} = \delta^2(\mathbf{q}_\perp - \mathbf{q}'_\perp)\delta(q - q') \tag{126}$$

where the volume element is $\mathrm{d}\mathcal{V} = \mathrm{d}x\mathrm{d}y\mathrm{d}z/z^3$. This condition is equivalent to the orthonormality of the functions $(2\pi)^{-1}e^{i\mathbf{q}_\perp \cdot \mathbf{x}}$ plus the relation [73, 74, 75]

$$\frac{2\pi^{-2}}{q}\sinh(\pi q)\int_0^\infty K_{iq}(z)K_{iq'}(z)\mathrm{d}z/z = \delta(q - q') \tag{127}$$

The expansion of a generic perturbation [73] in terms of these functions is

$$f(\mathbf{x}, z) = \int_0^\infty \mathrm{d}q \int \mathrm{d}^2\mathbf{q}_\perp f_{k\mathbf{q}_\perp}W_{k\mathbf{q}_\perp}(\mathbf{x}, z) \tag{128}$$

Following Wilson [14], we construct a Gaussian perturbation by assigning independent Gaussian distributions to the real and imaginary parts of the coefficients. Translation invariance in the $\mathbf{x}$ plane requires that the real and imaginary part of each coefficient has the same mean square, so we define the spectrum by (cf. Eq. (23))

$$\langle f^*_{k\mathbf{q}_\perp}f_{k'\mathbf{q}'_\perp}\rangle = \frac{2\pi^2}{q(q^2 + 1)}\mathcal{P}_f(k)\delta(q - q')\delta^2(\mathbf{q}_\perp - \mathbf{q}'_\perp) \tag{129}$$

The correlation function is

$$\xi_f(\mathbf{x}, z, \mathbf{x}', z') = \int_0^\infty \mathrm{d}q\frac{2\pi^2 N^2(k)}{q(q^2 + 1)}\mathcal{P}_f(k)\int \mathrm{d}^2\mathbf{q}_\perp e^{i\mathbf{q}_\perp \cdot (\mathbf{x} - \mathbf{x}')}zK_{iq}(q_\perp z)z'K_{iq}(q_\perp z') \tag{130}$$

It is invariant under Eq. (121), because the factor C appearing in Eq. (121) can be absorbed into the definition of the integration variable $\mathbf{q}_\perp$.

Using Eq. (122) and integrating over the angular direction in the $\mathbf{q}_\perp$ plane, the correlation function becomes

$$\xi_f = 2\pi\beta\int_0^\infty \mathrm{d}q\frac{\sinh(\pi q)}{\pi^2(q^2 + 1)}\mathcal{P}_f(k)\int_0^\infty \mathrm{d}ppJ_0(p\gamma)K_{iq}(p)K_{iq}(\beta p) \tag{131}$$

where

$$\beta \equiv \frac{\cosh\eta}{\cosh(\eta - r)} \tag{132}$$

$$\gamma \equiv \frac{\sinh r}{\cosh(\eta - r)} \tag{133}$$

From Eq. (8.13.30) of [74],

$$\int_0^\infty \mathrm{d}ppJ_0(p\gamma)K_{iq}(p)K_{iq}(\beta p) = \frac{\sqrt{\pi}\Gamma(1 + iq)\Gamma(1 - iq)}{2^{2/3}\beta(u^2 - 1)^{1/4}}P_{iq-1/2}^{-1/2}(u) \tag{134}$$

where
$$u \equiv \frac{\gamma^2 + \beta^2 + 1}{2\beta} = \cosh r \tag{135}$$

This is a function only of r, and using Eq. (8) one finds that the correlation function is given by Eq. (27).

Now consider the super-curvature modes $-1 < q^2 < 0$. The normalisation factor $N(k)$ becomes purely imaginary so it is convenient to drop the i factor, defining

$$N^2(k) = \frac{|q| \sin(\pi|q|)}{2\pi^4} \tag{136}$$

The super-curvature contributions are

$$f^{SC}(\mathbf{x}, z) = \int_0^1 d(iq) \int d^2\mathbf{q}_\perp f_{k\mathbf{q}_\perp} W_{k\mathbf{q}_\perp}(\mathbf{x}, z) \tag{137}$$

$$\langle f^*_{k\mathbf{q}_\perp} f_{k'\mathbf{q}'_\perp} \rangle = \frac{2\pi^2}{|q|(q^2 + 1)} \mathcal{P}_f(k) \delta(iq - iq') \delta^2(\mathbf{q}_\perp - \mathbf{q}'_\perp) \tag{138}$$

The correlation function is given by Eq. (131) with $q \to iq$, and since Eq. (134) is valid for $-1 < iq < 1$ this proves Eq. (44).

The spherical and flat-surface expansions are equivalent, at least in the present context, because they both lead to a Gaussian perturbation with the same correlation function.

The flat-space limit of the flat-surface expansion

We end by looking at the flat-space limit of the flat-surface expansion. Though not strictly necessary for our purpose, it is extremely instructive, and does not seem to have been given before.

The coordinates x, y and z become Cartesian in a a small region around $x = y = 0$ and $z = 1$, and then q_x, q_y and $q_z \equiv \sqrt{q^2 - q_x^2 - q_y^2}$ are the would-be components of the vector in the Fourier expansion. But the expansion Eq. (128) does not restrict the range of q_x and q_y, so it will include both real and imaginary q_z. In other words it will include hyperbolic functions as well as circular ones. Evaluating the limiting behaviour of K_{iq} confirms this. One finds [75] for real q_z

$$K_{iq}(q_\perp z) \to A \sin(B + q_z \tilde{z}) \tag{139}$$

where $\tilde{z} \equiv z - 1$, $A \equiv \sqrt{2\pi} q_z^{-1/2} e^{-\pi q/2}$ and $B \equiv \frac{\pi}{4} + q \cosh^{-1}(q/q_\perp) - q_z$. For imaginary q_z one finds

$$K_{iq}(q_\perp z) \to C \exp(D - |q_z|\tilde{z}) \tag{140}$$

where $C \equiv (2|q_z|/\pi)^{1/2}$ and $D \equiv -q_\perp \sin^{-1}(q/|q_z|)$.

The second expression becomes infinite when $|q_z|\tilde{z} \to -\infty$, but even so a translation invariant correlation function will result when the two expressions are substituted into Eq. (131). This example serves to remind us that we should take nothing for granted when considering which modes are allowable.

Acknowledgements

References

[1] P. Coles and G. Ellis, Nature **370**, 609 (1994).

[2] A. Dekel, Ann. Rev. Astron. Astroph., **32**, 371 (1994).

[3] G. F. R. Ellis, D. H. Lyth and M. B. Mijic, Phys. Lett B **271**, 52 (1991).

[4] D. H. Lyth and E. D. Stewart, Phys Lett. B **252**, 336 (1993).

[5] G. F. Smoot *et. al.*, *Astrophys. J. Lett.* **396** (1992) L1.

[6] K. Górski *et al.*, Astroph. J. Lett. **430**, 89 (1994).

[7] E. F. Bunn, D. Scott and M. White, preprint astro-ph/9409003 (1994).

[8] V. Fock, Z. Physik **98**, 148 (1935).

[9] M. Bander and C. Itzykson, Rev. Mod. Phys. **38** 346 (1966).

[10] L. P. Grishchuk and Ya. B. Zel´dovich, Astron. Zh. **55**, 209 (1978) [Sov. Astron. **22**, 125 (1978)].

[11] A. M. Yaglom, in *Proceedings of the Fourth Berkeley Symposium Volume II*, edited by J. Neyman (University of California Press, Berkeley, 1961).

[12] M. G. Krein, Ukrain. Mat. Z. **1**, No. 1, 64 (1949); *ibid* **2**, No. 1, 10 (1950).

[13] E. M. Lifshitz, J. Phys. (Moscow) **10**, 116 (1946); E. M. Lifshitz and I. M. Khalatinikov, Adv. Phys. **12**, 185 (1963).

[14] M. L. Wilson, Astrophys. J. **273**, 2 (1983).

[15] E. W. Kolb and M. S. Turner, *The Early Universe* (Addison-Wesley 1990).

[16] A. Z. Dolginov and I. N. Toptygin, Soviet Physics JETP, **37(10)**, 1022 (1960).

[17] N. Ya. Vilenkin and Ya. A. Smorodinsky, Soviet Phys. JETP, **19**, 1209 (1964).

[18] E. R. Harrison, Rev. Mod. Phys. **39**, 862 (1967).

[19] N. D. Birrel and P. C. W. Davies, *Quantum Field Theory in Curved Space-Time* (Cambridge University Press 1982).

[20] R. Fabbri, I. Guidi and V. Natale, Astrophys. J. **257** 17 (1982).

[21] I. M. Gelfand, M. I. Graev and N Ya Vilenkin, *Generalized Functions: Volume 5; Integral Geometry and Representation Theory* (Academic Press Inc., New York, 1966); M. A. Naimark, *Linear Representations of the Lorentz Group* (Pergamon Press 1964).

[22] R. J. Adler, *The Geometry of Random Fields* (Wiley, Chichester, 1981).

[23] P. J. E. Peebles, *The Large Scale Structure of the Universe* (Princeton University Press, 1980).

[24] A. D. Linde, *Phys Lett* **B129**, 177 (1983).

[25] S. Coleman and F. De Luccia, Phys. Rev. D **21**, 3305 (1980); J. R. Gott, Nature **295** , 304 (1982); A. H. Guth and E. J. Weinberg, Nucl. Phys. **B212**, 321 (1983); J. R. Gott and T. S. Statler, Phys. Lett. **B136**, 157 (1984).

[26] M. Bucher, A. Goldhaber and N. Turok, preprint (1994).

[27] M. Sasaki, T. Tanaka, K. Yamamoto and J. Yokoyama, Phys. Lett. B **317**, 510 (1993); M. Sasaki, T. Tanaka, K. Yamamoto and J. Yokoyama, Prog. Theor. Phys. **90**, 1019 (1993); T. Tanaka and M. Sasaki, preprint (1994). T. Tanaka and M. Sasaki, Phys. Rev. D , in press (1994); K. Yamamoto, T. Tanaka and M. Sasaki, preprint (1994).

[28] M. Bucher and N. Turok, in preparation.

[29] S. Karlin and H. M. Taylor, *A first course on stochastic processes* (Academic Press, New York, 1975).

[30] J. M. Bardeen, J. R. Bond, N. Kaiser and A. S. Szalay, *Astrophys. J.* **304**, 15 (1986).

[31] M. L. Wilson, Astrophys. J. Lett., **253**, 53 (1982).

[32] K. M. Gorski and J. Silk, Astrophys. J. **346**, L1 (1989).

[33] N. Sugiyama, N. Gouda and M. Sasaki, Astrophys. J. **365**, 432 (1990).

[34] N. Gouda, N. Sugiyama and M. Sasaki, Prog. Theor. Phys., **85**, 1023 (1991).

[35] M. Kamionkowski and D. N Spergel, Astrophys. J. **432**, 7 (1994).

[36] N. Sugiyama and J. Silk, Phys. Rev. Lett. **73**, 509 (1994).

[37] M. Kamionkowski, D. N. Spergel and N. Sugiyama, Astrophys. J. **426** (1994) L57.

[38] M. Kamionkowski, B. Ratra, D. N. Spergel and N. Sugiyama, "CBR anisotropy in an open inflation, CDM cosmogony", Princeton preprint (1994).

[39] B. Ratra and P. J. E. Peebles, Astroph. J., **432**, 5 (1994); B. Ratra and P. J. E. Peebles, preprint (1994).

[40] M. Tegmark and J. Silk, "Reionization in an open CDM universe: implications for cosmic microwave background fluctuations", Berkeley preprint (1994).

[41] A. Kashlinsky, I. Tkachev and J. Friedman, Phys. Rev. Lett., **73**, 1582 (1994).

[42] J. M. Bardeen, Phys. Rev. D, **22**, 1882 (1980).

[43] D. H. Lyth, Phys. Rev. D **31**, 1792 (1985).

[44] D. H. Lyth and M. Mukherjee, Phys. Rev. D **38**, 485 (1988); D. H. Lyth and E. D. Stewart, Astrophys. J. **361**, 343 (1990).

[45] G. F. R. Ellis and M. Bruni, Phys. Rev. D **40**, 1804 (1989); M. Bruni, P. K. S. Dunsby and G. F. R. Ellis, Astrophys. J. **395**, 34 (1992); P. K. S. Dunsby, M. Bruni and G. F. R. Ellis, Astrophys. J. **395**, 54 (1992).

[46] A. R. Liddle and D. H. Lyth, Phys. Rep. **231**, 1 (1993).

[47] A. R. Liddle and D. H. Lyth, Mon. Not. R. Astron. Soc., to be published (1994); A. R. Liddle and D. H. Lyth, in preparation.

[48] H. Kodama and M. Sasaki, Prog. Theor. Phys. **78**, 1 (1984).

[49] V.F. Mukhanov, H. A. Feldman and R. H. Brandenberger, Phys. Rep. **215**, 203 (1992).

[50] D. S. Salopek, J. R. Bond and J. M. Bardeen, Phys. Rev. **D40**, 1753 (1989).

[51] J. M. Bardeen, P. S. Steinhardt and M. S. Turner, Phys. Rev. D **28**, 679 (1983).

[52] D. H. Lyth and E. D. Stewart, in preparation; M. Bruni and P. K. Dunsby, Int. Journ. Mod. Phys **3**, 443 (1994).

[53] D. H. Lyth and E. D. Stewart, Astrophys. Journ. **361**, 343 (1990).

[54] M. Bruni and D. H. Lyth, Phys. Lett. B **323**, 118 (1994).

[55] A. A. Starobinsky, Phys. Lett. **B117**, 175 (1982); S. W. Hawking, Phys. Lett. **B115**, 339 (1982); A. H. Guth and S.-Y. Pi, Phys. Rev. Lett. **49**, 1110 (1982).

[56] V. F. Mukhanov, JETP Lett. **41**, 493 (1985).

[57] M. Sasaki, Prog. Theor. Phys. **76**, 1086 (1986). Phys. Lett. B **323**, 118 (1994).

[58] E. J. Copeland, A. R. Liddle, D. H. Lyth, E. D. Stewart and D. Wands, Phys. Rev. D **49**, (1993), page 6427.

[59] D. H. Lyth, Phys Letts **147B**, 403 (1984).

[60] J. J. Halliwell and S. W. Hawking, Phys. Rev. D **31**, 1777 (1985).

[61] A. D. Linde, Mod. Phys. Lett, **A1**, 81 (1986); A. D. Linde, *Particle Physics and Inflationary Cosmology* (Harwood, Chur, Switzerland, 1990).

[62] A. D. Linde, 'Lectures on Inflationary Cosmology' unpublished (1994).

[63] E. D. Stewart, personal communication.

[64] E. D. Stewart, preprint hep-ph/940539, to be published in Phys. Rev. D; E. D. Stewart, preprint astro-ph/9407040, to be published in Phys. Letts. B; E. D. Stewart, preprint hep-ph/9408302; D. H. Lyth and E. D. Stewart, preprint hep-ph/9408342, to be published in the proceedings of 'Birth of the Universe and Fundamental Physics'; G. Dvali, Q. Shafi and R. Schaefer, preprint hep-ph/9406139 (1994).

[65] E. D. Stewart and D. H. Lyth, Phys. Lett. B **302**, 171 (1993).

[66] A. A. Starobinsky, Sov. Astron. Lett. **11**, 133 (1985).

[67] B. Allen and R. Caldwell, preprint (1994).

[68] R. Wald, *Quantum Field Theory in Curved Spacetime* (University of Chicago Press 1994).

[69] K. M. Gorski, H. Ratra, N. Sugiyama and A. J. Banday, preprint (1994).

[70] R. R. Caldwell and A. Stebbins, preprint (1994).

[71] M. S. Turner, Phys Rev D **44**, 12 (1991).

[72] M. Sasaki, T. Tanaka and K. Yamamoto, preprint (1994).

[73] N. N. Lebedev, C. R. (Doklady) Acad. Sci. URSS (N.S.) **58**, 1007 (1946).

[74] A. Erdelyi, Ed., *Tables of Integral Transforms*, (McGraw-Hill, 1953).

[75] A. Erdelyi, Ed., *Higher Transcendental Functions, Volume II* (McGraw Hill, 1953).